ELECTRONIC SWITCHING TECHNIQUES FOR FIBER OPTIC NETWORKS

by

Hyun Jong Shin

Memorandum No. UCB/ERL M88/15

23 February 1988

ELECTRONIC SWITCHING TECHNIQUES FOR FIBER OPTIC NETWORKS

by

Hyun Jong Shin

Memorandum No. UCB/ERL M88/15

23 February 1988

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ELECTRONIC SWITCHING TECHNIQUES FOR FIBER OPTIC NETWORKS

by

Hyun Jong Shin

Memorandum No. UCB/ERL M88/15

23 February 1988

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ELECTRONIC SWITCHING TECHNIQUES FOR FIBER OPTIC NETWORKS

Ph.D. Hyun Jong Shin Department of EECS

David G. Hodges
Chairman of Committee

ABSTRACT

There is a growing demand for wideband circuit switching networks to provide non-blocking channels for broadband communication with mixed services including voice, data, and video. In this dissertation, electronic switching techniques for high data rate, star-topology, optical fiber networks are studied. This thesis emphasizes the integrated circuit implementation of a crosspoint switch, which is the key element in circuit switching, with a capacity of 100 to 200 Mb/s.

CMOS technologies with minimum channel lengths below 3 μm have been chosen to permit data propagation at the desired speeds, while minimizing power consumption and chip area. Circuit techniques that may reduce timing jitter and noise arising from capacitive and inductive couplings between on-chip nodes including power supply rails have been explored. One effective technique is to limit the voltage swing at critical nodes.

As a demonstration vehicle, a prototype switch containing crosspoints for 16 inputs and 16 outputs as well as associated control circuitry has been designed and fabricated. Proper operation has been obtained with the samples of three different minimum channel lengths: 2.8, 2, and 1.4 μm. This shows the feasibility that prevalent, inexpensive, and relatively low-speed CMOS VLSI technologies are applicable to 100 to 250 Mb/s circuit switching networks.

Measurements have been carried out with a 2^{15}-1 sequence, pseudo-random, non-return-to-zero data. Crosstalk between channels has been observed by sending data to one output while broadcasting other asynchronous data to the other 15 outputs. The switch with the 1.4 μm gate length operates up to 250 Mb/s with 80% eye opening and 0.2 ns jitter from a 5 V supply. The worst crosstalk adds 140 mV peak-to-peak noise. Average delay through the switch is 5.9 ns and power dissipation is 900 mW. Performances of 2 μm and 2.8 μm chips scale according to the processes: 215 Mb/s and 105 Mb/s with 80% eye opening, respectively. The outputs can drive other identical circuits to enable construction of larger switching networks. External transistors are necessary to drive the 150 mA required by typical laser diodes.

To my parents, wife Hye Kyung, and sons Woo Chul and Yoon Chul

for their love and encouragement

ACKNOWLEDGEMENT

The author would like to express his hearty appreciation to Professor David A. Hodges for the thoughtful guidance and support throughout the course of this work. The author also wishes to thank Professor David G. Messerschmitt and Professor Paul R. Gray for their valuable advice and encouragement.

Discussions with Dr. Ming-Kang Liu who looked at system aspects of fiber optic networks during his Ph.D. work were indispensable. Fellow students including Sehat Sutarja, Bosco Leung, Wen-Bin Hsu, Joey Doernberg, and Steve Lewis were helpful in many respects.

Special thanks go to Dr. George Lewicki and staffs at MOSIS, the General Electric Company, and Dr. Ping Ko, Ms. Katalin Voros, and staff at the U.C. Berkeley Microfabrication Laboratory for the fabrication support.

This research was supported in part by grants from Advanced Micro Devices, Bell Communications Research, Fairchild Semiconductor, Harris Corporation, Intel, National Semiconductor, and Xerox, with matching grants from the State of Californa MICRO program, and in part by the National Science Foundation under grant number MIP-8603430.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Optical fiber that carries information in the form of light is an ideal transmission medium in wideband communication networks because of the unrivalled high bandwidth and low attenuation characteristics: bit rate over 1 Gb/s, loss as low as 0.2 dB/km, and resulting bandwidth-distance product on the order of 100 Gb/s × km [1]. As well, it is more secure from tapping and immune to electromagnetic interference and crosstalk compared to other transmission links.

Along with this fact, the recent development of fiber optics that has resulted in cheap, reliable, and easily installable fibers has stimulated a substantial interest in fiber based networks and opened a wide variety of possibilities in future communications. Examples include fiber optic long-distance network [2], fiber *local area network* (LAN) [3-7], broadband *integrated services digital network* (ISDN) [8-11], and optical interconnects [12].

Most of these applications focus on broadband digital communication and services. Especially in the area of fiber optic LAN and broadband ISDN, important new services include video exchange and distribution. To facilitate this kind of services, there is a growing demand for broadband circuit switching that provides high data rate non-blocking communication channels. Broadband switching necessitates, first of all, the development of wideband circuit switching elements such as multiple channel crosspoint switches, although the determination of network structure, medium access protocol, and service integration method is equally important.

In order to preserve the inherent advantages of optical communication, primarily the high bandwidth, it is better to process information in optical domain. For broadband switching, this would be realized by implementing switches using optical techniques. However, optical

techniques are not mature yet and pose interfacing problems with surrounding electronics. On the other hand, electronic techniques are becoming viable for wideband applications, because of the dramatic improvement in speed recently driven by *very-large-scale integration* (VLSI) and *very-high-speed integrated-circuit* (VHSIC) technologies.

In this dissertation, studies of electronic switching techniques for high data rate, optical fiber networks are described. The objective of this work is to apply electronic techniques to broadband, star-topology, circuit switching networks. Emphasis is put on the *integrated-circuit* (IC) implementation of the key element in circuit switching, a crosspoint switch, with a bit rate adequate for digital video. More specifically, this thesis investigates whether a multiple channel crosspoint switch with a capacity of 100 to 200 Mb/s is feasible in widely available CMOS VLSI technologies.

CHAPTER 2

FIBER OPTIC CIRCUIT SWITCHING NETWORKS

Demand for broadband LANs and ISDNs that feature integration of real-time and interactive video services is beginning to increase because information is delivered more effectively in visual form. Fiber networks with circuit switching would meet this demand due to unique advantages over others without switching. In this chapter, the advantages of circuit switching in fiber optic networks are examined. Also, some issues related to switching networks such as access method and network expansion will be briefly discussed.

2.1 FIBER OPTIC NON-SWITCHING NETWORKS

As for any network technology, the utilization efficiency of a fiber network largely depends on its topology and medium access method [13]. Most common topologies for fiber networks are star and ring [7,14]. Figure 2.1 shows non-switching star and ring networks.

A star topology fiber network combined with the *carrier-sense multiple-access / collision-detection* (CSMA/CD) protocol results in network performance like that for Ethernet [15], but at higher data rate. These LANs are classified as passive or active depending on whether the center of a network consists of passive optical power couplers or active devices that amplify optical signals.

Passive star LANs, although very simple, may be troublesome because transmitted power is partially lost in the star coupler and divided by the number of stations [3]. Thus, a passive star is not adequate for a network with many stations. Also, it is hard to expand the network due to the

difficulty of optically coupling extra fiber links onto an existing star coupler. For passive star, the multiple access control function or the collision detection and resolution function is typically distributed to the network interface at each station.

On the other hand, in active star networks, each network transmitter needs transmission power only to reach the center. Expansion is easier since every new station requires only a point-to-point bidirectional fiber link to the center. However, minimum signal power requirements at the center limit the network size and expansion. Contention resolution is normally performed at the center [4] or hub [5] in an active star network.

A ring-topology network with token-passing protocol constitutes a LAN that has orderly access control and higher network efficiency [6]. This network architecture is suitable for higher bit rate applications. But reliability of the network is poor when a single ring is used. Any breakage of the ring causes complete shutdown of the network. Like star networks, a ring network can be passive or active. An active ring is advantageous for expansion because the network interface does not require lossy passive couplers and, hence, addition of a fiber link between stations is easy [7].

The networks discussed above can be classified as broadcasting networks in the sense that any source station transmits information to all stations in the network regardless of actual destinations. Therefore, energy launched into the network has to be higher than that required for the desired communication and most of it is dissipated in vain. Additionally, in general, only one source is given the right to occupy the network at any instant and others have to wait until the network becomes free. Once it is free, awaiting stations try to grab the network and contention may occur between them. This contention is then resolved by the access protocols. A broadcasting network efficiently utilizes the network bandwidth when the traffic is light. However, it cannot effectively support long and heavy traffic.

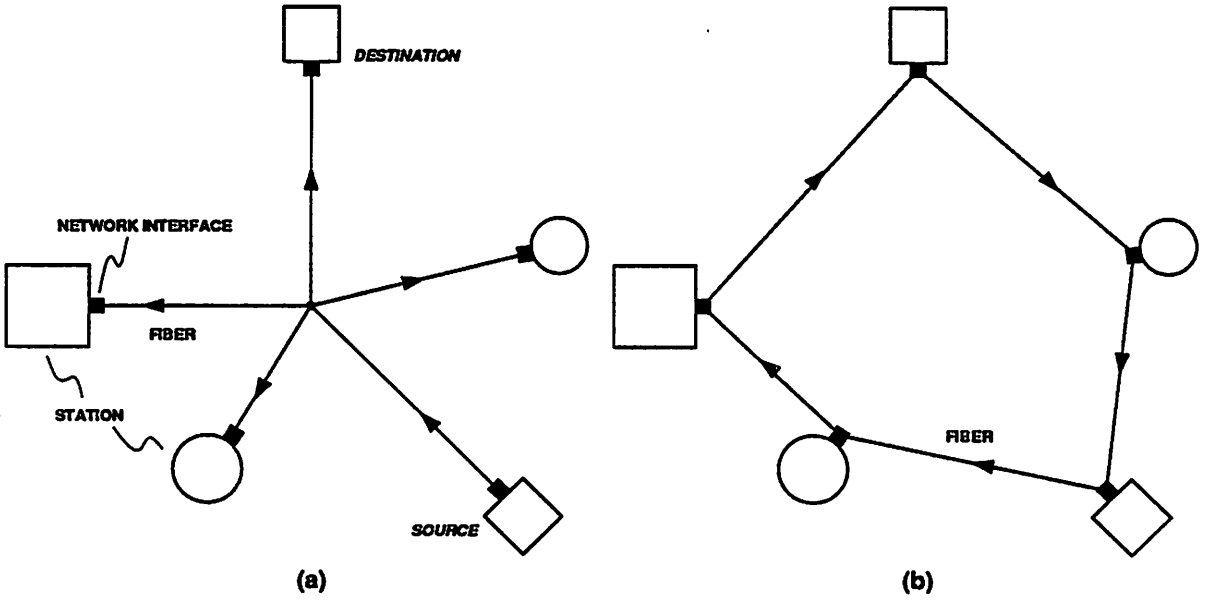


Figure 2.1 (a) Star- or (b) ring-topology fiber optic local area networks without circuit switching.

2.2 FIBER OPTIC SWITCHING NETWORKS

To allow uninterrupted, real-time traffic, a network needs capability to establish multiple concurrent communication channels. One way to achieve this ability is to install fiber links in every possible combination between the stations, as shown in Figure 2.2 (a), and let each station to set up paths to its own destinations. However, this method is impractical because the number of links required is excessive if the number of stations is large; for N stations, $(N-1)!$ bidirectional links are necessary.

Figure 2.2 (b) shows another solution that involves a switching function at a common node or center, and fibers between this node and the stations. This is a natural extension of an active star network and is logically equivalent to the above network except that the distributed switching is centralized. This network is easily expandable because the number of fiber links is equal to that of stations.

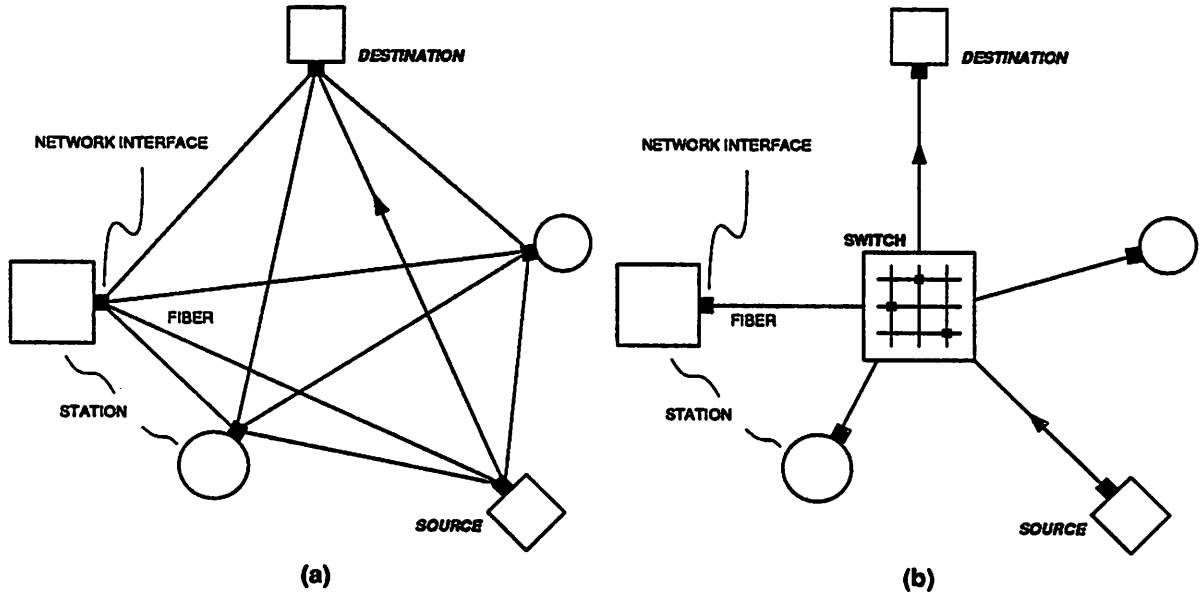


Figure 2.2 (a) Distributed or (b) centralized switching fiber optic networks.

The switching mechanism at the center of a star network can be packet switching or circuit switching [16]. Packet switching relies on high-speed processing of packet headers that contain switching information to route packets. It shows higher network efficiency, because traffic on all channels has less idle time. When a channel is occupied by heavy traffic, newly arrived packets are temporarily stored in memory and later, when the channel becomes idle, the stored packets are transmitted. If overall traffic is light, the delay in transmission is minor. However, if it is heavy, serious delays (*latency*) may occur. In the latter case, storage requirements will be increased as well. Due to this latency and requirement for high-speed devices, packet switching is not attractive for real-time wideband applications. Large latency times are not acceptable in two-way voice or video communication services.

Circuit switching incorporated in a broadband star network sets up combinations of concurrent physical circuits for data transmission. These circuits are transparent in the sense that any input will appear at an output without being processed or delayed by more than the physical transmission time. Hence, circuit switching is attractive for high data rate, real-time

communication.

2.3 CIRCUIT-SWITCHING FIBER NETWORKS

2.3.1 Advantages

As illustrated in Figure 2.3, circuit switching allows simultaneous, blocking-free data exchanges including both conversation and broadcasting. This feature is the primary advantage of circuit switching over the non-switching networks [17].

The improvement can be noted first by considering the maximum possible number of concurrent communication channels. When every source is connected to a unique destination, the number reaches a maximum that is equal to the number of stations. Obviously the minimum occurs when a station broadcasts to all destinations, like the non-switching network. Secondly, note the number of connection combinations between stations at any time. For N stations, the circuit-switching networks have $(N + 1)^N$ possible combinations while the non-switching ones have only $N + 1$.

Other advantages are related to signal energy, network bandwidth, and versatility. Because communication paths are established only between active stations, unnecessary energy loss is minimized in circuit-switching networks. Also, due to the transparency of the switch mentioned in the previous section, higher network bandwidth can be obtained. The switch permits configuration of various, topologically equivalent networks, such as bus, ring, or non-switching star. In addition, circuit switching eases network management and enhances reliability. However, it requires expensive, wideband switching components and effective ways to access them.

2.3.2 Application

One application area of circuit switching would be fiber LANs. Figure 2.4 depicts an example star-topology fiber LAN with circuit switching. In this LAN, information exchanges are possible

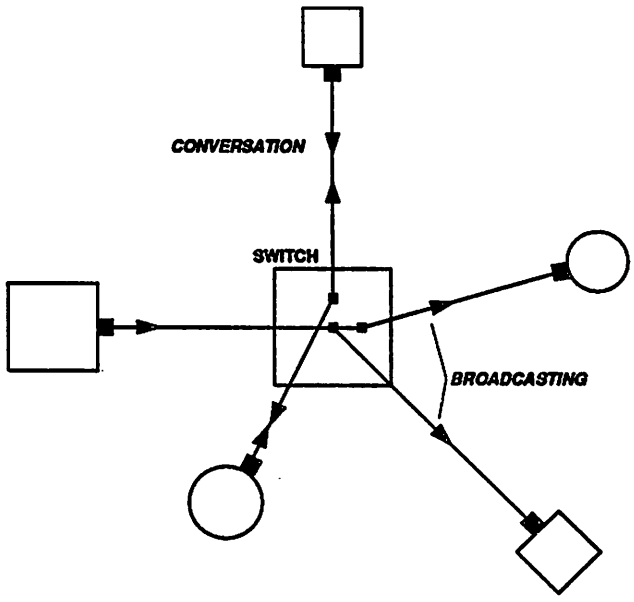


Figure 2.3 Concurrent traffic in circuit-switching fiber networks, supporting broadcasting and conversation.

in large quantities, at high data rates, between computer subsystems (processors, storage devices, and terminals) through the switch. The simultaneously available, parallel channels alleviate the bottleneck of a single bus, allow direct transfer between devices, and increase total system throughput. The non-blocking nature of this network facilitates special features like selective broadcasting of messages, real-time video or graphics data transfer, and video conferencing.

Another area that demands wideband circuit switching is future broadband ISDNs [8-11]. A broadband ISDN can be envisioned as in Figure 2.5, where optical fiber is used in subscriber loops and digital video, audio, and video-phone services are combined with narrowband telephone and data services. In this network, separate circuit switches are provided at the local exchange center for different services, so that each subscriber can select various video or audio sources and communicate with other subscribers via phones or data links independently, without blocking.

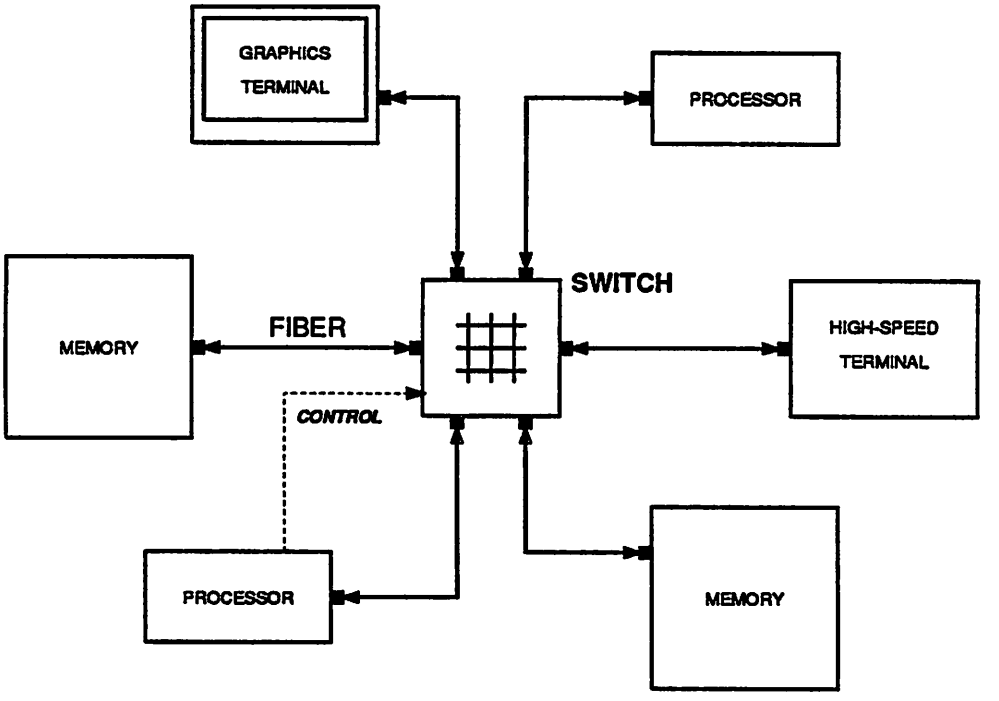


Figure 2.4 A star-topology, circuit-switching, fiber-optic LAN.

2.3.3 Crosspoint Switch

Obviously, one primary element for circuit switching is the multiple channel crosspoint switch. Figure 2.6 shows a non-blocking crosspoint switch for N stations that has $N \times N$ switch cells or crosspoints: one crosspoint at every intersection of input and output links. Logically, the $N \times N$ switch is a set of N, N-input multiplexers. In this basic switch, each switch cell is open or closed depending on switching information supplied from an external controller. One restriction on the setup of the switch is that, among N cells connected to any output link, only one cell can be closed at any time. This restriction prevents collision of data between two or more sources and allows only one source to occupy the channel to any destination. But it does not limit the number of output links connectable to one input link.

Bandwidth requirement on the switch is determined by the maximum bit rate of data in a network. In the two kinds of broadband switching networks discussed in the previous subsection,

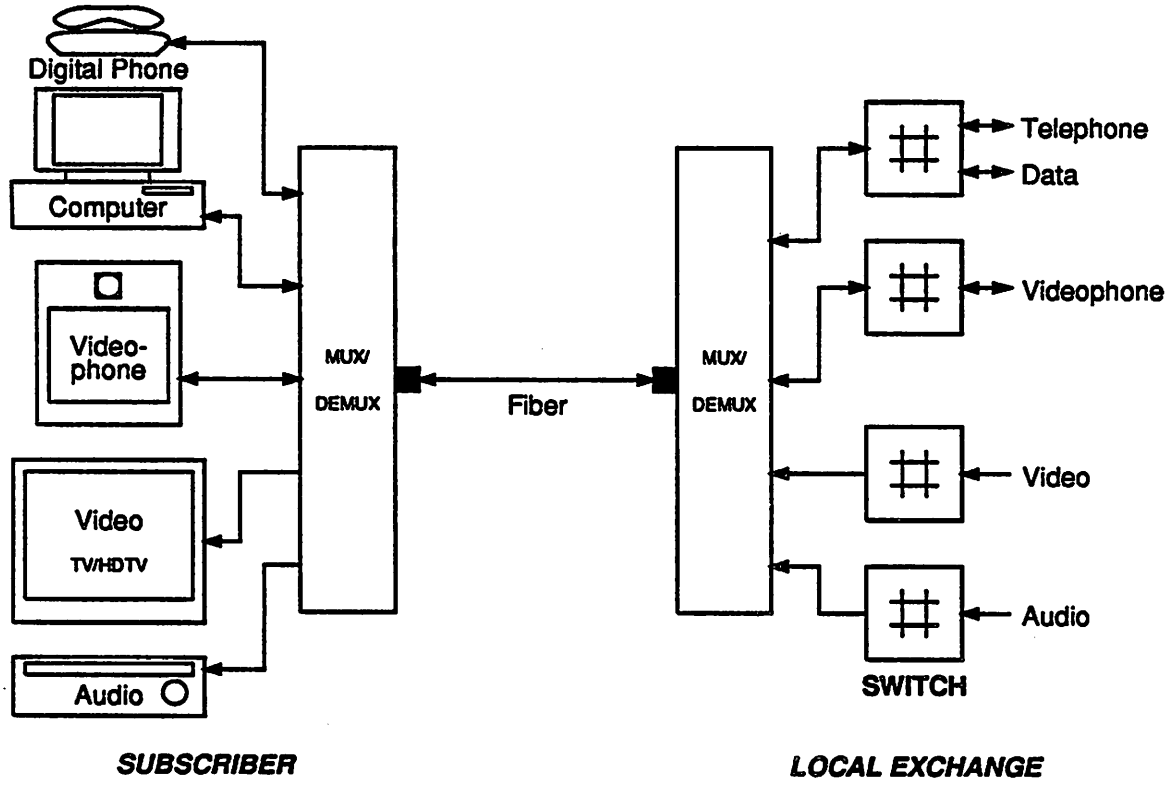


Figure 2.5 Future broadband integrated services digital network.

the switch bandwidth is mainly determined by the digitized video signals whose bit rates are as high as 140 Mb/s [9-11]. However, the switch bandwidth is not equal to the control rate or switch setup rate. The control rate can be substantially lower than the bandwidth, although it depends on the requests from every station and varies a lot with applications.

2.3.4 Medium Access Methods

As in a telephone network, medium access in a circuit-switching fiber network is a procedure to set up circuits or physical channels for intended communication. Communication circuits are established through the switches according to the switching information. To access the medium, each station must request connection changes by calling a switch controller. Depending on the characteristics of the communication service, either sources or destinations may request changes.

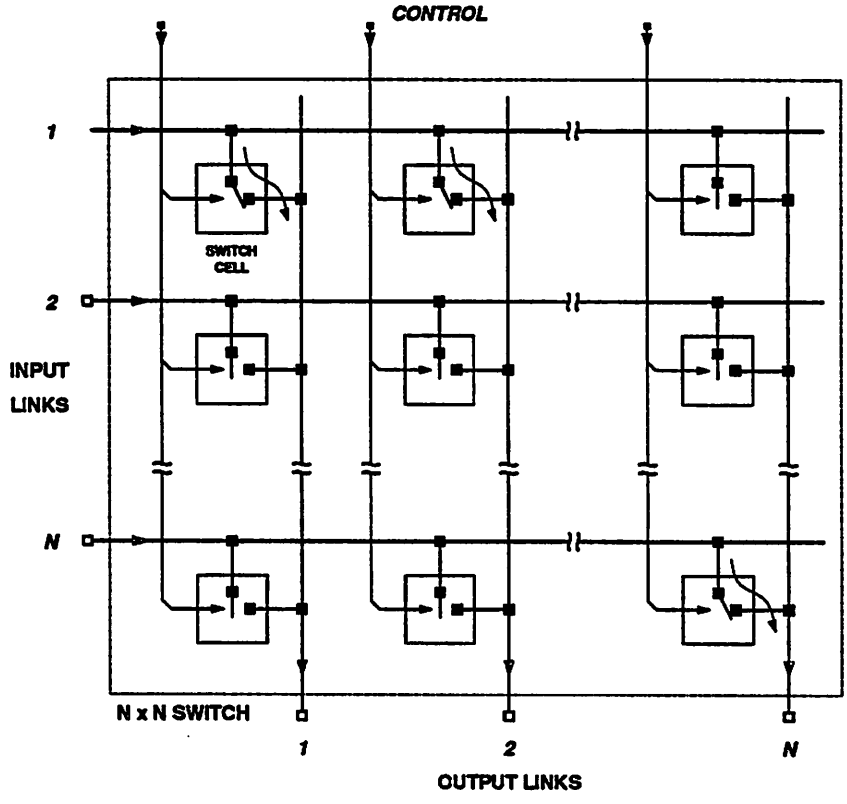


Figure 2.6 Non-blocking N x N crosspoint switch.

The requests can be made immediately before the transmission, or long before it in case of pre-scheduling. After receiving the requests, the switch controller resolves conflicts among requests and grants the medium access rights to proper stations.

To facilitate the request/grant process without interrupting the main wideband communication, separate channels must be provided for switch control. One method to provide switch control channels is depicted in Figure 2.7. In this method, a physically separate, separate network carries the calls. Because the requesting rate is much lower than the data rate for most applications, a narrowband copper network such as Ethernet would be sufficient. Also a dedicated wire pair can be utilized if a station is not linked to a call network. A host station controls the switch, and receives and acknowledges requests through the call channels. Although this configuration allows flexibility in handling high-speed data and calls, it complicates networking. However,

efforts to add the call network may not be necessary when an existing narrowband LAN is upgraded to include broadband services by putting a fiber switching network, backbone.

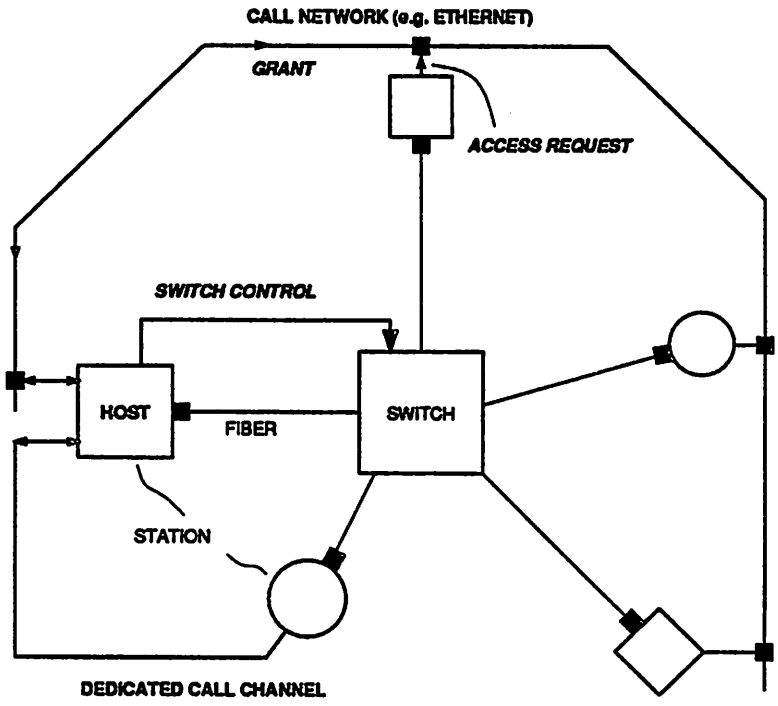


Figure 2.7 Medium access in a circuit-switching fiber network through a physically separate low-speed call network.

In another network shown in Figure 2.8, switch control information can share the same transmission medium, i.e., fiber, with the main communication. Hence, multiplexing of two different kinds of information is needed in this method. One promising multiplexing technique is the wave division multiplexing where light of different wavelengths carries different information that can be separated by corresponding optical filters. Another technique is to add a header in front of the data that contains the switching information. The header would have a lower bit rate to ease its processing. This technique is similar to packet switching in a sense, but the header may contain future schedules as well as requests that must be executed immediately.

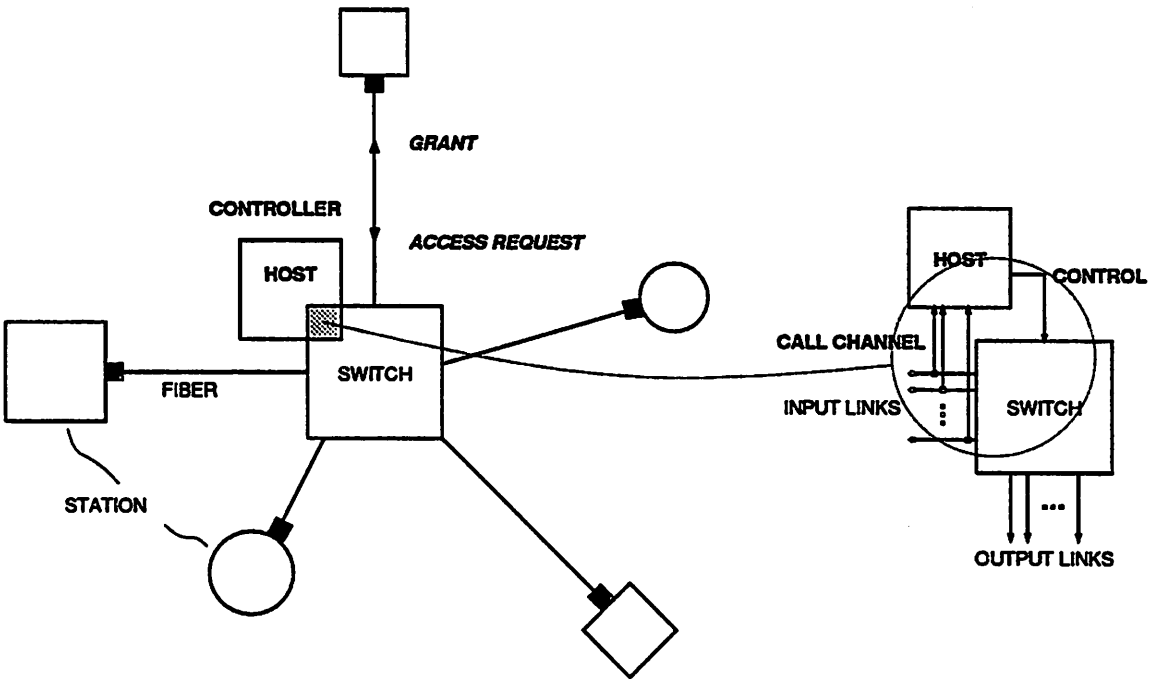


Figure 2.8 Medium access in a circuit-switching fiber network through medium sharing.

2.3.5 Expansion

A way to expand a switching network or to enlarge the switch size is to connect an array of switches as shown in Figure 2.9. In this array, input and output of every switch are busses connected in parallel to the input links and output links, respectively. Each individual switch needs an open-circuit state output when no input is connected. This array is non-blocking and is equivalent to one large size crosspoint switch. However, to realize a $M \times M$ network using switches of $N \times N$ crosspoints, $[M/N]^2$ switches are needed, where $[M/N]$ represents the integer between M/N and $M/N + 1$ if M/N is not an integer, or M/N, otherwise. For example, a 1024-link network will require 4096 16-channel switches. Another drawback of this array is that an active output draws more power to drive the capacitive open-circuit outputs as well.

The output power burden can be reduced by using a switch that has bypass connections [9,28,29]. Figure 2.10 shows the switch and a non-blocking array composed of these switches.

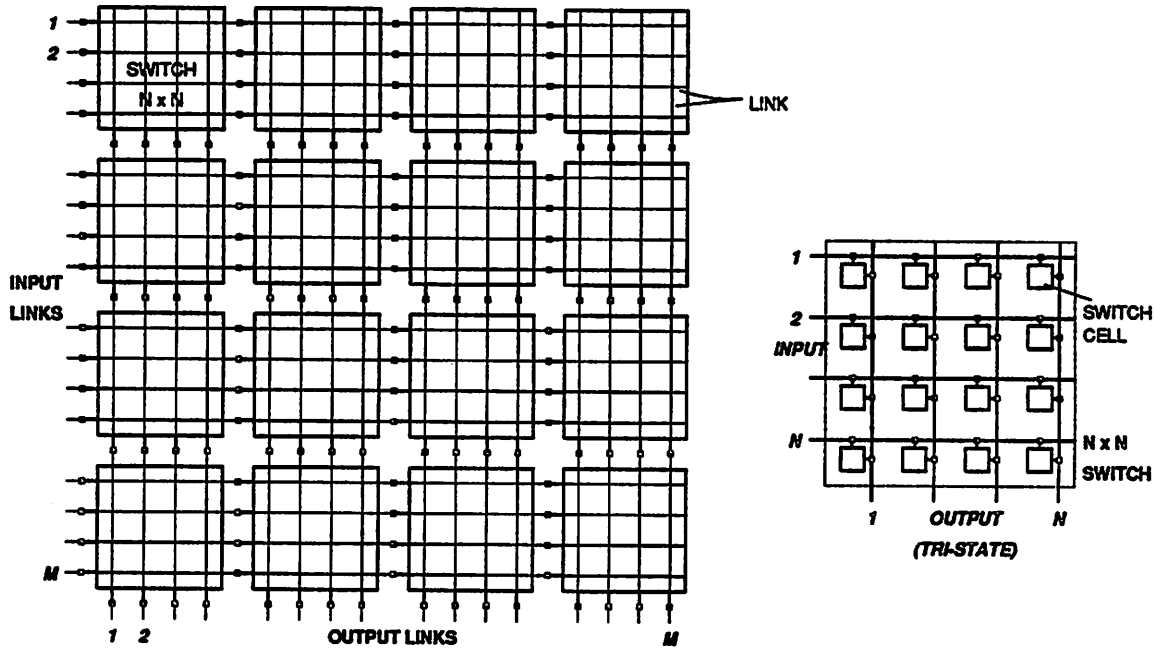


Figure 2.9 Non-blocking expansion of a switching fiber network using switches with N × N crosspoints.

The switch has N more switch cells in addition to the ordinary $N \times N$ crosspoints. These extra crosspoints provide vertical paths from the bypass or extension inputs to the outputs when the main inputs are not switched on. Like the previous array, the main inputs of each switch in this array are connected in parallel to the input links. However, the switches are serially cascaded in vertical direction, such that the bypass inputs of every switch are connected to the outputs of the upper switch. Therefore, each output needs to drive only one bypass input if the switch is an active device. An obvious disadvantage of this method is the increased complexity and number of input/output ports of the switch.

The previous switch expansion methods require a large number of switches, although they are non-blocking. For applications in which traffic is light among distant stations, the non-blocking feature for those channels can be sacrificed. In that case, an alternative method that connects switches in hierarchical structure such as a tree (as in conventional telephone switching net-

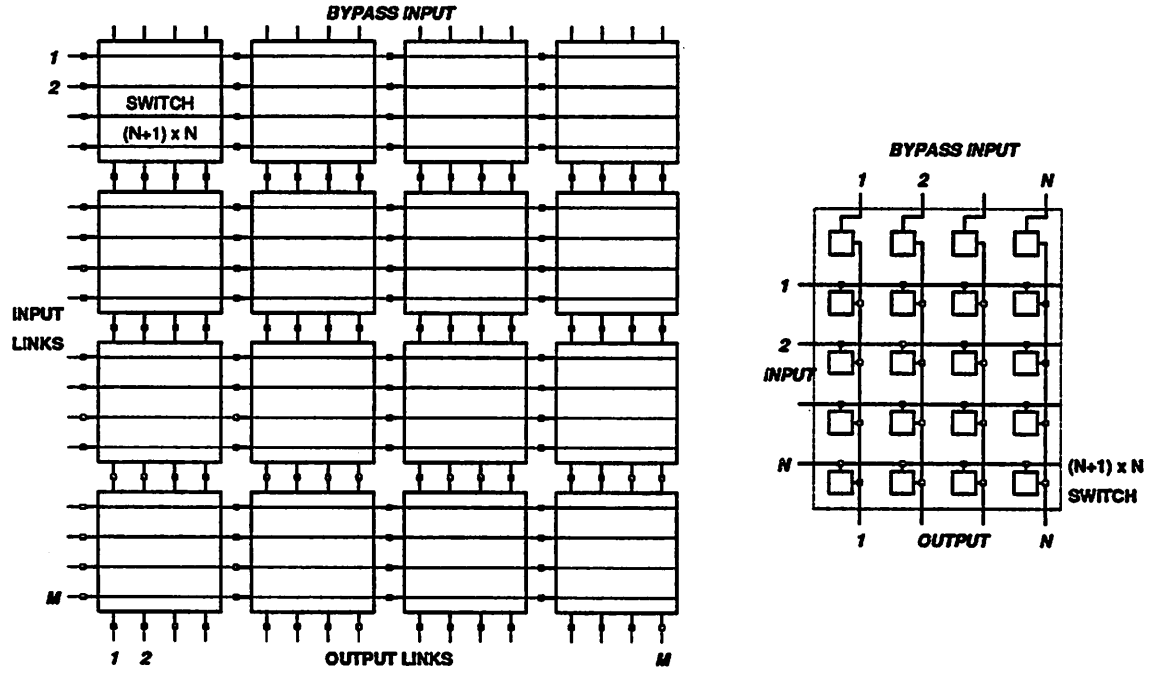


Figure 2.10 Non-blocking expansion of a switching fiber network using switches with (N+1) × N crosspoints.

works) would be sufficient. This concept is depicted in Figure 2.11, where transmission links between different hierarchical levels become bottlenecks and cause a serious blocking when interlevel traffic is heavy. But the blocking can be relieved by assigning multiple links between switches. Although the number of switches increases as blocking is minimized, this method drastically reduces the total number of switches. If single link is assigned for interlevel connection, the number becomes

$$[M/(N-1)] + [[M/(N-1)]/(N-1)] + [[[M/(N-1)]/(N-1)]/(N-1)] + \cdots$$

where the last term is 1. For the same 1024-link example with $N = 16$, only 75 switches and three levels of hierarchy are enough; the first level needs 69 switches, the second level 5, and the third level 1.

Another alternative to construct larger switching networks is to use the well-known 3-stage

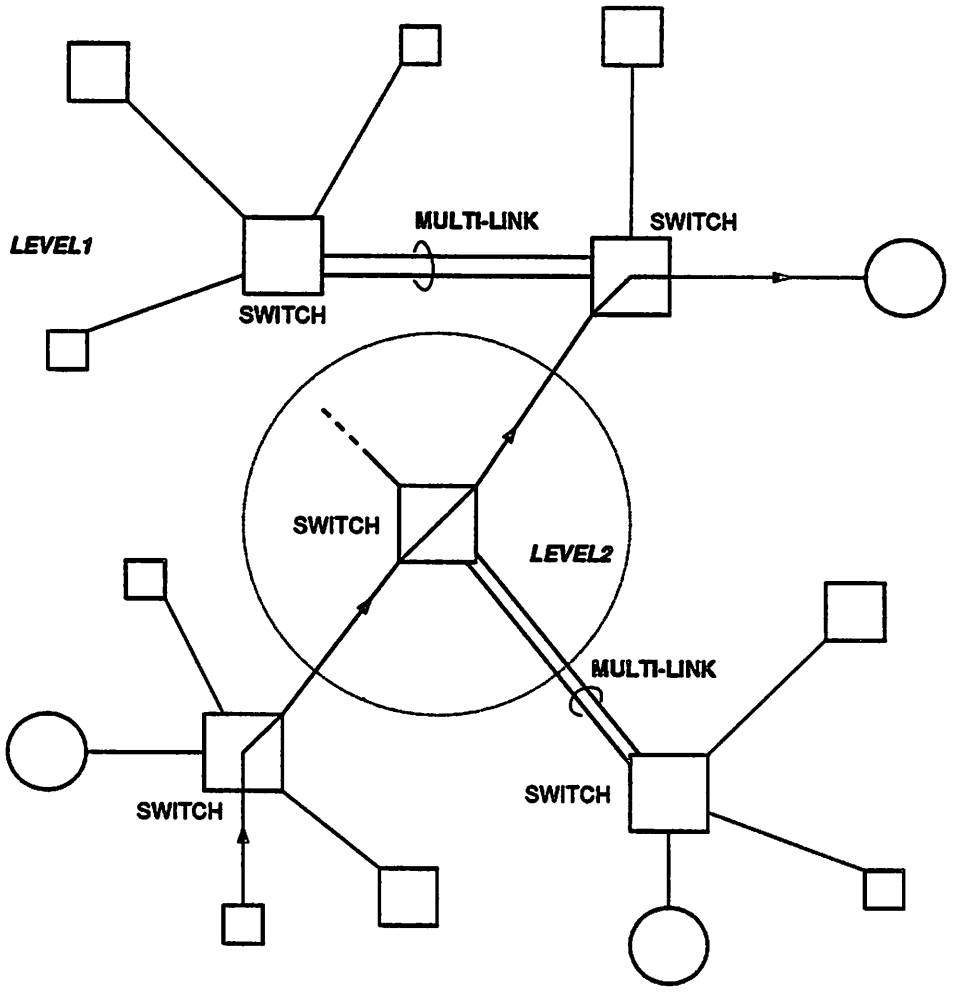


Figure 2.11 Expansion of a switching fiber network using hierarchical structure.

switch network [16,18,29] shown in Figure 2.12. This switch network is advantageous in that it achieves strict non-blocking with a reasonably small number of switches. Furthermore, the blocking is controllable by adjusting the number of switches or K. For $K = 1$, the resulting network is identical to the tree network discussed above. If K is increased, the number of the second stage switches increases, while the blocking decreases. Eventually, when $K = 2N - 1$, the network becomes non-blocking. One thing to note in this configuration is that switches of different crosspoint sizes are needed: $N \times K$, $[M/N] \times [M/N]$, and $K \times N$. These switches could be constructed using methods shown in Figures 2.9, 2.10, or even Figure 2.12 out of $N \times N$ switches.

To implement a $M \times M$ non-blocking switch in this structure with individual switches that are made using the method in Figure 2.9, $4[M/N] + (2N - 1)[[M/N]/N]^2$ switches of $N \times N$ crosspoints are required. For the same example ($M = 1000$ and $N = 16$), the total number of switches is 752.

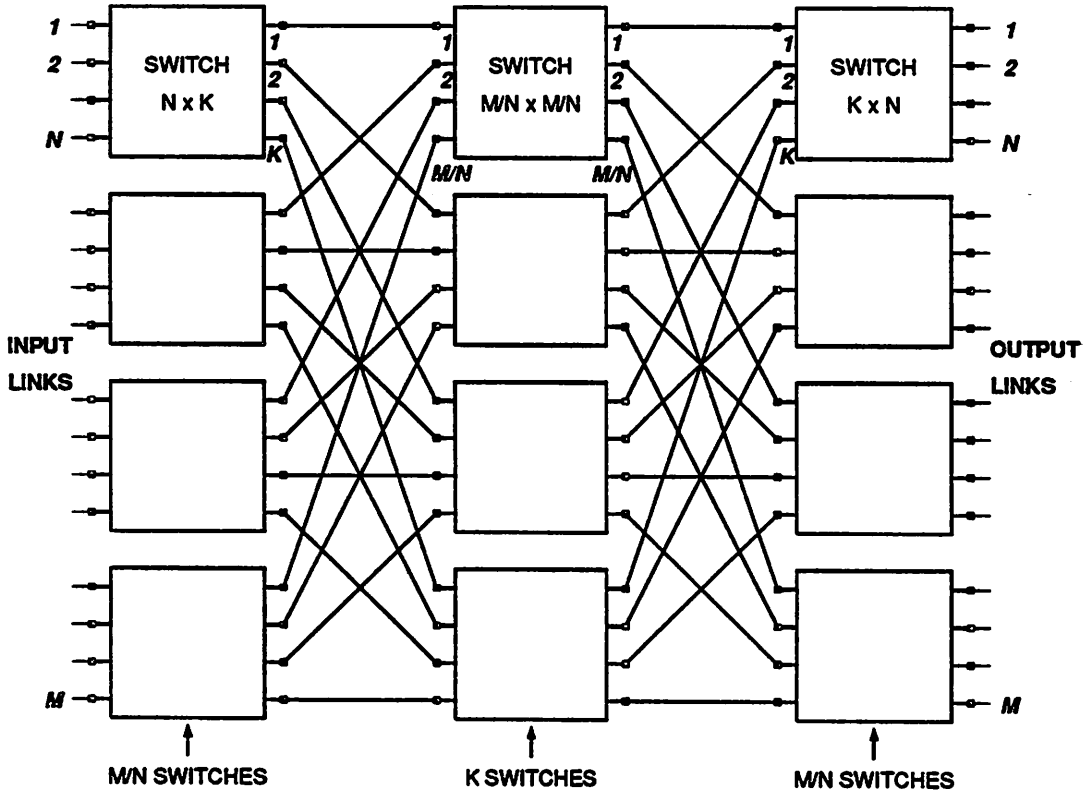


Figure 2.12 Expansion of a switching fiber network using 3-stage switch structure.

The square matrix structure and 3-stage structure are compared in Figure 2.13 for the number of $N \times N$ switches needed to build a non-blocking switch network for M stations; in this case, $N = 16$ and M varies up to 200. Also shown is the curve for a blocking, tree structure network with single interlevel links. The 3-stage network is more promising as M gets larger than 96. Below this, the square matrix results in less switch counts. Note that, for the 3-stage network when M is smaller than 129, the 16 × 16 switches are shared to construct the second stage switches.

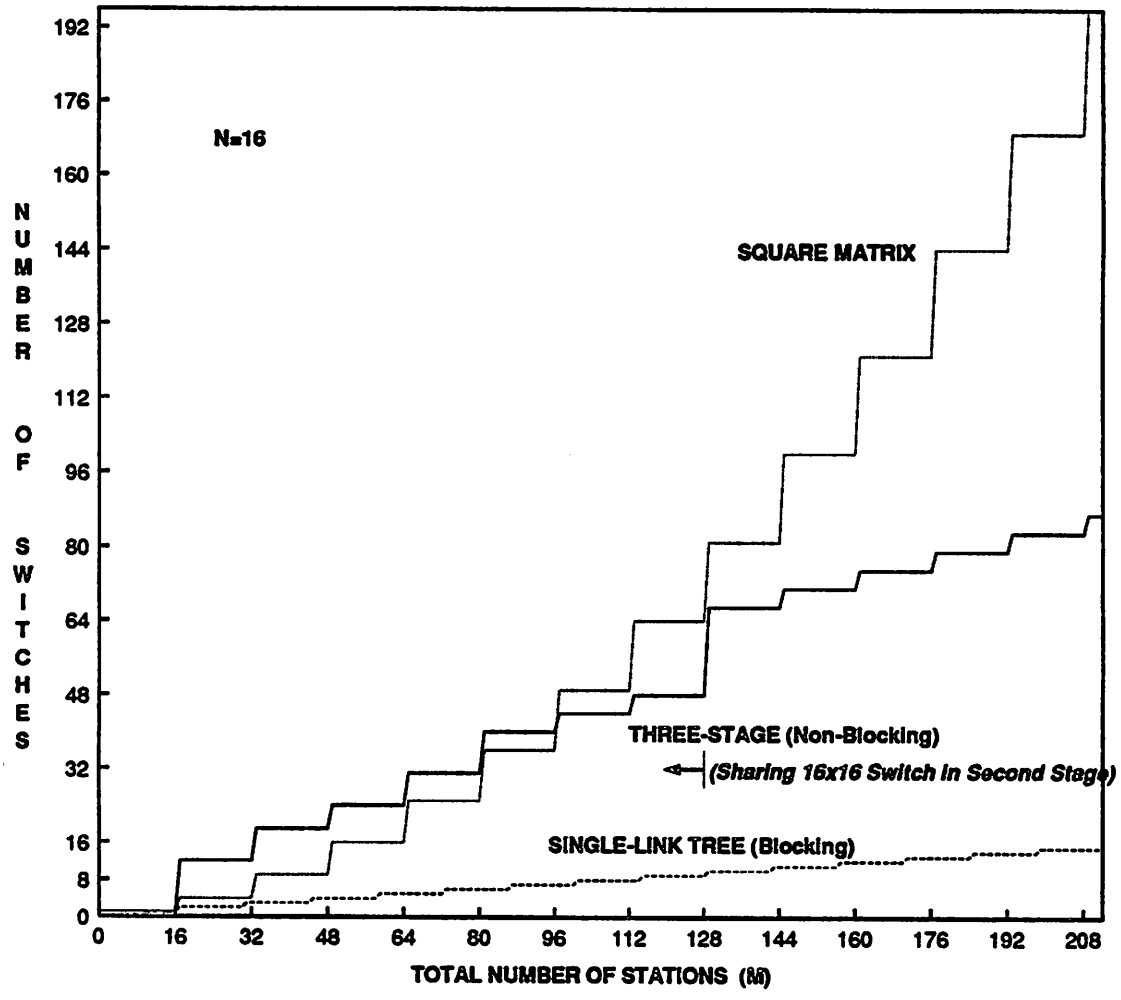


Figure 2.13 Comparison of the number of switches for various expansion methods versus number of stations.

CHAPTER 3

ELECTRONIC SWITCHING TECHNIQUES

As has been pointed out in Chapter 2, the primary element for circuit switching is the crosspoint switch. To be useful, the switch not only has to have the required bandwidth but also needs to support as many channels as possible and interface well with peripheral devices. Among these desired features, the bandwidth requirement is most important in selecting technologies for the switch. In general, technologies for implementing the switch can be classified as optical and electronic. According to this classification, broadband circuit switching techniques can also be categorized as optical and electronic.

In the first part of this chapter, electronic switching techniques are compared to optical switching techniques and their pros and cons are briefly addressed. Then, the rest part focuses on the issues of electronic switching techniques: technology choice for the required switch bandwidth (100 to 200 Mb/s) and switch architectures.

3.1 OPTICAL SWITCHING TECHNIQUES

To preserve the inherent advantages of optical communication, mainly the high bandwidth, one approach is to process information in a network entirely in optical form. In an optical network, any switching element should route the signal as desired with low energy loss. Therefore, in principle, an optical signal transmitted from a source through a fiber can be switched to another fiber with minimal waveform degradation.

Due to this prospective advantage, various optical switching techniques have been investi-

gated extensively. Among them are mechanical-optic [19], magneto-optic [20], and electro-optic techniques [21-25]. The main objectives of this prior work have been to reduce the crosstalk between channels, insertion loss, operating voltage and power, and bulkiness, for the construction of large switch arrays with good interface. The most promising techniques are electro-optic switches fabricated in integrated optics technologies, including the directional coupler [22] and waveguide crossover [24] switches shown in Figure 3.1. In the directional coupler switch, when no voltage is applied, light in a surface waveguide is fully transferred to the adjacent waveguide through the coupler. On application of a voltage, the coupling characteristic changes so that light is confined to the original waveguide. The waveguide crossover switch operates similar way, but the refractive index changes at the crossover and incident light is refracted to the other waveguide when a voltage is applied.

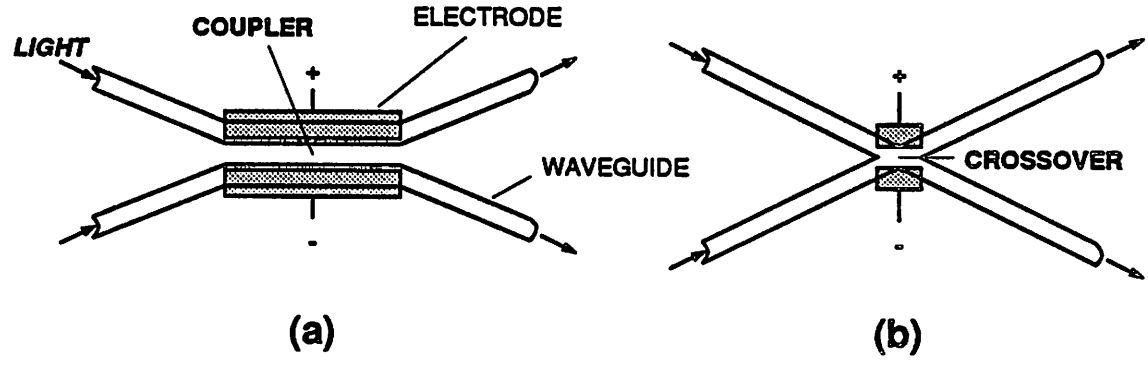


Figure 3.1 (a) Directional coupler or (b) waveguide crossover electro-optic switch in integrated optics technology.

However, the switches demonstrated so far suffer from poor characteristics: the crosstalk around -20 dB, insertion loss 1 dB/cm, size near 1 cm, and operating voltage from 10 to 20 V. Especially, the mechanical- and magneto-optic switches have been too bulky. Furthermore, it has not been easy to make reasonable size arrays with non-blocking and broadcasting capability out of these devices [21,25]. Another thing to note is that these optical switches are controlled by applying electrical power, i.e., voltage or current. Hence, presently available optical techniques

require an interface to surrounding electronics. In other words, the application of optical techniques is limited only to the high speed channel, and the control should be realized in electronic technologies.

3.2 ADVANTAGES AND DISADVANTAGES

Even though electronic technologies are still inferior in bandwidth to optical ones, there has been a dramatic increase in speed with the rapid development of semiconductor VLSI and VHSIC technologies [26,27]. As a result, electronic techniques have become viable for certain wideband applications, such as point-to-point optical fiber communication up to GHz range [28-35] and broadband (about 140 Mb/s) circuit switching networks [36-40].

Electronic switching techniques based on the state-of-the-art IC technologies are advantageous over optical ones in many respects. They are more reliable and cheap, provide better interface to peripheral electronics, and allow construction of a miniature switch with a reasonable number of crosspoints. Also, integration of control functions in a switch is possible. Furthermore, since the devices used are mostly active, specification of the loss characteristic that is important for passive devices is not needed. Another advantage to note is that the crosstalk can be minimized by applying proper techniques in the design.

However, electronic techniques necessitate a number of interface devices in the high speed transmission channel that perform electrical-optical and optical-electrical signal conversion with enough bandwidth and gain. Figure 3.2 illustrates that basically optical receivers and transmitters are needed to support electronic switching in a fiber optic network. In its simplest form, an optical receiver is composed of a photodetector and a wideband amplifier, while an optical transmitter consists of a current driver and a laser diode or light-emitting diode. Note that the switch may be a single device or a module expanded using one of the methods discussed in Section 2.3.5.

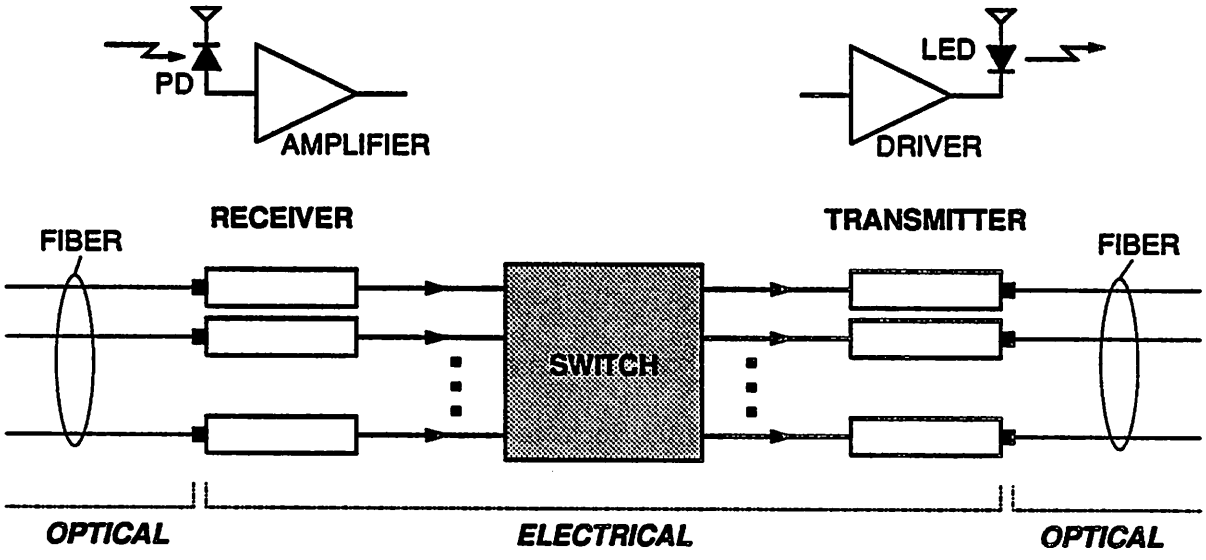


Figure 3.2 Electronic switching technique in a fiber optic network: optical receivers and transmitters are needed.

One important drawback of electronic broadband switching is that, since the bandwidth of a technology used is not sufficiently larger than the data rate, the components cause pulse dispersion and degrade timing information. Also electromagnetic interference and crosscoupling impose noise on the pulses. In short, data waveforms in electronic switching are more likely distorted with timing jitter and noise as illustrated in Figure 3.3. In other words, the eye pattern that is a superposition of real data waveforms in every ideal bit interval gets more closed. Here, timing jitter is defined as a variation of time points extracted from the pulse transitions crossing a reference level, relative to the original clock edges.

3.3 TECHNOLOGY CHOICE

3.3.1 Data Transmission and Coding

In digital data communication, information is contained in a sequence of highs and lows and the associated pulse transitions. To retrieve information, in general, transmitted data must be sam-

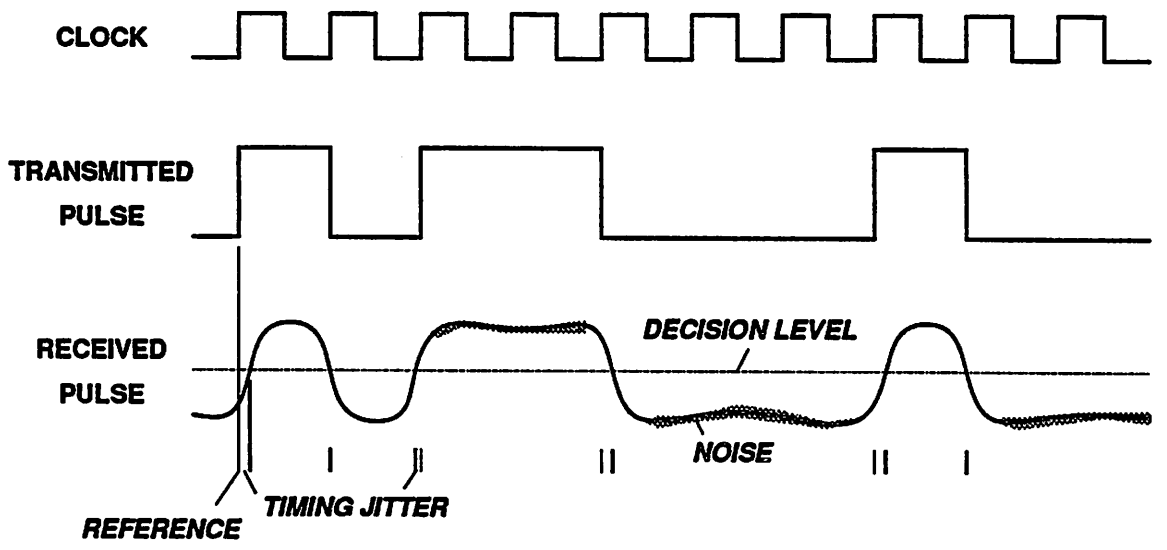


Figure 3.3 Timing jitter and noise on received pulses.

pled at proper time points with a clock that can be derived from the pulse transitions. For a well-defined clock, the high/low levels must be adequate to reduce transmission errors. Therefore, to minimize errors in transmission in a switching network, the pulse distortion should be restricted to a certain degree. This means that the eye opening must be sufficient.

Normally, *phase-locked loops* (PLLs) are used to extract timing information from the data and generate a jitter free clock whose frequency tracks the data rate. However, functioning of a PLL is assured when there are enough pulse transitions and the amount of timing jitter is not excessive.

The number of pulse transitions per unit time interval depends on the coding methods used to transmit data. In the popular *non-return-to-zero* (NRZ) binary coding, the data are encoded such that '1' and '0' plainly correspond to high and low levels, respectively. Because the bit rate of actual transmission is equal to the data rate, the NRZ coding requires a minimum channel bandwidth. But, since it may have no pulse transition for a long period of time, it would make timing extraction difficult. An improvement in this respect is the *return-to-zero* (RZ) coding that

encodes '1' into a pulse returning to the low level. Timing extraction gets easier, but still there may be no pulse for a substantial period. In addition, the channel bandwidth requirement doubles. The Manchester coding that converts '1' into high-low transition and '0' into low-high ensures one transition in every bit period. Although clock can be easily extracted from this code, it also needs a channel bandwidth twice the data rate. Intermediate between the NRZ and Manchester coding, there are other coding methods which maintain certain transition densities at the expense of bit rates. One example is the 4B/5B coding [6] that takes 4-bit data and encodes into 5-bit symbols. This achieves at least one pulse transition within 3 bit periods, with a 25% higher bit rate as an exchange. Another example is the fixed transition block coding [17].

The pulse degradation expected from an inferior technology can partly be reduced by designing components such that the technology bandwidth is maximally utilized. This may involve circuit techniques that would improve the switch bandwidth but sacrifice other performance parameters instead. If the improvement is not sufficient, at some places on the communication channels, pulses must be re-shaped or regenerated before the distortion reaches a critical value. In other words, pulse regeneration devices are needed which restore correct levels and timing. Of course, regeneration devices consisting of timing extraction, decision, and synchronization circuits must be implemented with higher bandwidth technologies.

3.3.2 Bit Rate

Appropriate technologies for the implementation of switches should at first satisfy the bandwidth requirement. Among prospective technologies that meet the requirement, however, it is better to choose ones that are cheap, reliable, and suitable for large scale integration or consume less power and area.

The required bandwidth is directly related to the maximum bit rate of transmission through the switch. And, as mentioned briefly, the bit rate depends on the encoding schemes. For exam-

ple, for 140 Mb/s digital video services, the bit rate would vary from 140 to 280 Mb/s. If the 4B/5B coding is used, actual bit rate may go up to 175 Mb/s. Because NRZ or other coding schemes with low transition densities are likely to be used for efficient utilization of channel bandwidth, to support 140 Mb/s digital video, a switch bandwidth of 200 Mb/s would be sufficient. This is 100 Mb/s for 70 Mb/s video. Obviously, the switch would have enough bandwidth margin for these coding methods, in case it is designed to accommodate the maximum bit rate.

3.3.3 Selection Criteria

Let us consider a gate driven by a f_C Hz, 50% duty-cycle clock that is equivalent to a $2f_C$ b/s NRZ signal when the transition density is maximum (Figure 3.4). Also, let us assume that the gate delay mainly comes from charging the capacitive output and the charging current I is constant. In this case, the capacitance C_L represents either another gate 4 times larger in size or a load equivalent to that. This is assumed because proceeding from inputs to outputs where large drives are required, for optimum delay and driving capability, sizes of successive gates need to increase by a certain factor. For MOS technologies, this factor is about 4 (Appendix 1).

Because the combination of gate and load has a limited slew rate, for ideal input transitions, the output will rise and fall with finite slopes. To ensure a proper signal transfer, however, the rise and fall transitions must be completed within times shorter than one bit period. Since the rise time (t_R) and fall time (t_F) are defined between 10% and 90% of the full swing $V_H - V_L$ and the transitions are linear, the above condition can be restated as

$$\frac{5}{4} max[\, t_R, t_F \,] \leq \frac{1}{2} t_C. \tag{3.1}$$

Assuming symmetric transitions,

$$t_R = t_F \leq 0.4 t_C. \tag{3.2}$$

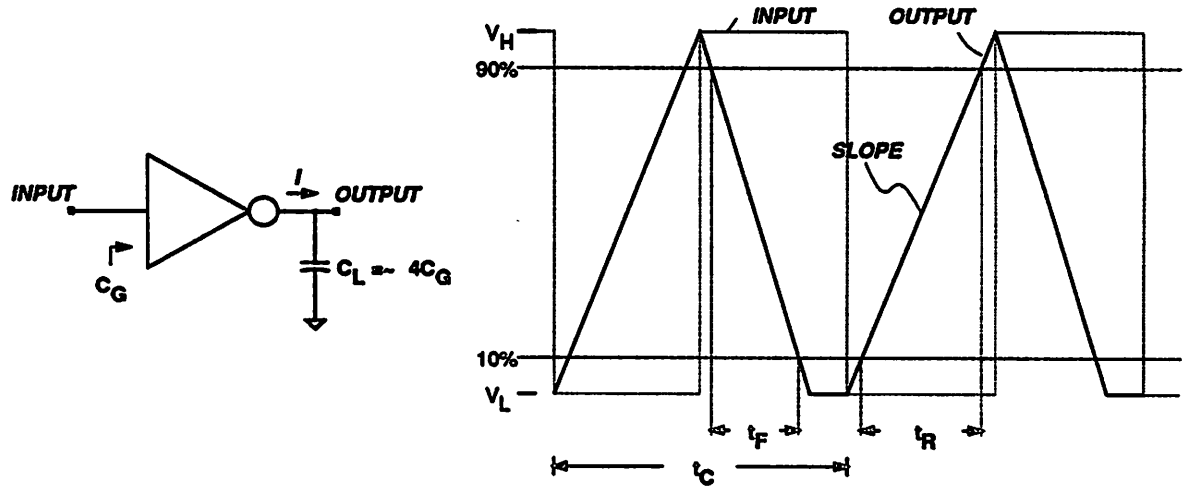


Figure 3.4 A gate driving a capacitive load.

Because the technology bandwidth is conventionally specified as the gate delay for fan-out of 1, it is better to state the above relationship in terms of this. Now, note that the delay time is approximately $\frac{5}{8}$ of the rise time and, if the parasitics at the gate output are ignored, the rise time for fan-out of 1 is roughly $\frac{1}{4}\,t_R$. Therefore, the gate delay, τ, must satisfy

$$\tau \leq \frac{1}{16}\,t_C. \tag{3.3}$$

As an example, for t_C = 10 ns (f_C = 100 MHz), the gate delay must be smaller than 0.625 ns. If the optimum fan-out is greater than 4, the upper limit would decrease. But, as shown in Appendix 1, a larger optimum fan-out presumes more parasitics at the gate output. Using the relationship in Eq.(A1.3), the upper limit increases due to the higher parasitics.

The technology bandwidth is also represented as the unity gain frequency. Supposing that the circuit is not slew limited, adequate information is propagated if, at least, the fundamental frequency component of the clock passes through with a gain not smaller than one. This means that the unity gain bandwidth for fan-out of 4 should be larger than the clock frequency. In other words, the unity gain frequency, f_T, defined for fan-out = 1 must satisfy

$$f_T \geq 4f_C, \tag{3.4}$$

when the output parasitics are neglected. For a 100 MHz clock, this must be greater than 400 MHz.

The Eqs. (3.3) and (3.4) are selection criteria for a technology. Between the two, however, Eq. (3.3) is dominating for both bipolar and MOS technologies. In typical bipolar circuits,

$$f_T = \frac{g_m}{2\pi C_G} \tag{3.5}$$

and

$$\tau = \frac{C_G(V_H - V_L)}{2I}. \tag{3.6}$$

Using $g_m = \frac{qI}{kT}$, the ratio between f_T and the slew limiting frequency, $\frac{1}{4\tau}$, is

$$4\tau f_T = \frac{V_H - V_L}{\pi \frac{kT}{q}}. \tag{3.7}$$

For $(V_H - V_L) = 1$ V, the ratio is about 12. Because the swing is on the order of 1 V or larger, the unity gain frequency is much higher than the slew limiting frequency. The above equations for f_T and τ are also valid for MOS circuits. However, in long-channel MOS circuits, g_m is expressed as $\frac{2I}{(V_{GS} - V_T)}$ where V_{GS} is the gate bias and V_T is the transistor threshold voltage. Thus, the ratio becomes

$$4\tau f_T = \frac{2(V_H - V_L)}{\pi(V_{GS} - V_T)}. \tag{3.8}$$

For a 5 V swing and 1 V threshold, if the circuit is biased at the center of the swing, the ratio is around 2. The ratio is larger than this value in cases where the bias voltage is much closer to V_T.

3.3.4 Technology Choices

Technologies that allow implementation of 200 Mb/s switches and fulfill the condition (3.3) and (3.4) include modern GaAs, silicon bipolar, and MOS technologies. GaAs technologies have

been among the fastest and 1 μm gate length devices show τ near 100 ps [27] and f_T of 7.5 GHz [35]. Silicon bipolar technologies have become competitive in speed with GaAs and, for emitter stripe widths below 3 μm, τ is less than 500 ps [26] and f_T is over 6 GHz [28-30]. MOS technologies, especially NMOS, with effective channel lengths below 1 μm have demonstrated τ less than 70 ps [31]. Typical MOS gate delays are about 400 ps for 2 μm drawn channel lengths and 200 ps for 1.5 μm [26].

Among these, CMOS technologies with channel lengths below 2 μm may be most advantageous for this application because of their suitability for VLSI in terms of static power consumption and area. For the implementation of crosspoints only, there is nothing special about CMOS technologies that are still inferior in speed and drive capability. But a switch would need integration of control functions or processing units and memories to be easily interfaceable with other system components. This prospective integration level excludes other technologies dissipating too much power and area. In fact, for the above reasons, CMOS technologies have been selected in most of the switches reported so far [36,38-40], although it has been possible to reduce power dissipation in bipolar switches [37]. Note also that the recent drive toward VLSI has spurred enormous development and proliferation of CMOS technologies, making them cheap and reliable.

3.4 SWITCH ARCHITECTURES

The multiple channel crosspoint switch can be structured as either matrix or multiple multiplexer form. Also the switch may be constructed differently depending on whether the switch setup is requested primarily from sources or destinations. Source driven setup would be more useful for LAN applications while destination driven one would be better for broadband ISDN applications.

3.4.1 Multiplexer Architecture

As illustrated in Figure 3.5, the multiplexer architecture literally consists of multiplexers, one for every output link. Each multiplexer requires a separate register that stores the source address to be selected and a decoder that decodes this address. A $M \times N$ switch needs N copies of M-to-1 multiplexers. The bit size of each register is normally $[\ln_2 M]$. If it is allowed not to select any of the sources, M is replaced by $M+1$. This architecture is suitable for destination driven setup requests. A link is set up by selecting the source address register according to the destination address and writing the register. Major problems of this structure include a large number of devices and interconnects for the decoders that makes the layout difficult and inefficient.

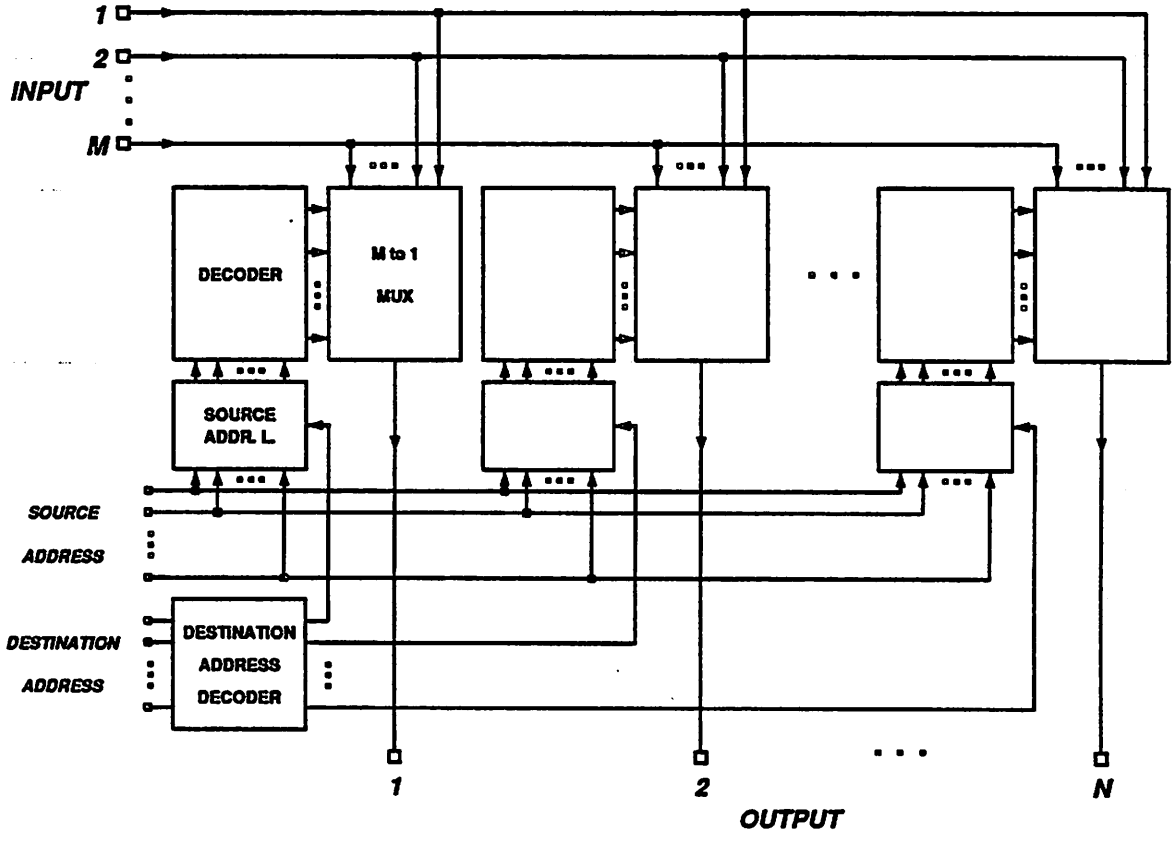


Figure 3.5 Multiplexer type switch architecture.

3.4.2 Matrix Architecture

A layout efficient architecture for a switch is of the matrix type. In this architecture, a switch cell is placed between the input and output at every intersection and a single bit memory is connected to this cell. For a $M \times N$ switch, $M \times N$ pairs of switch and memory cell are required.

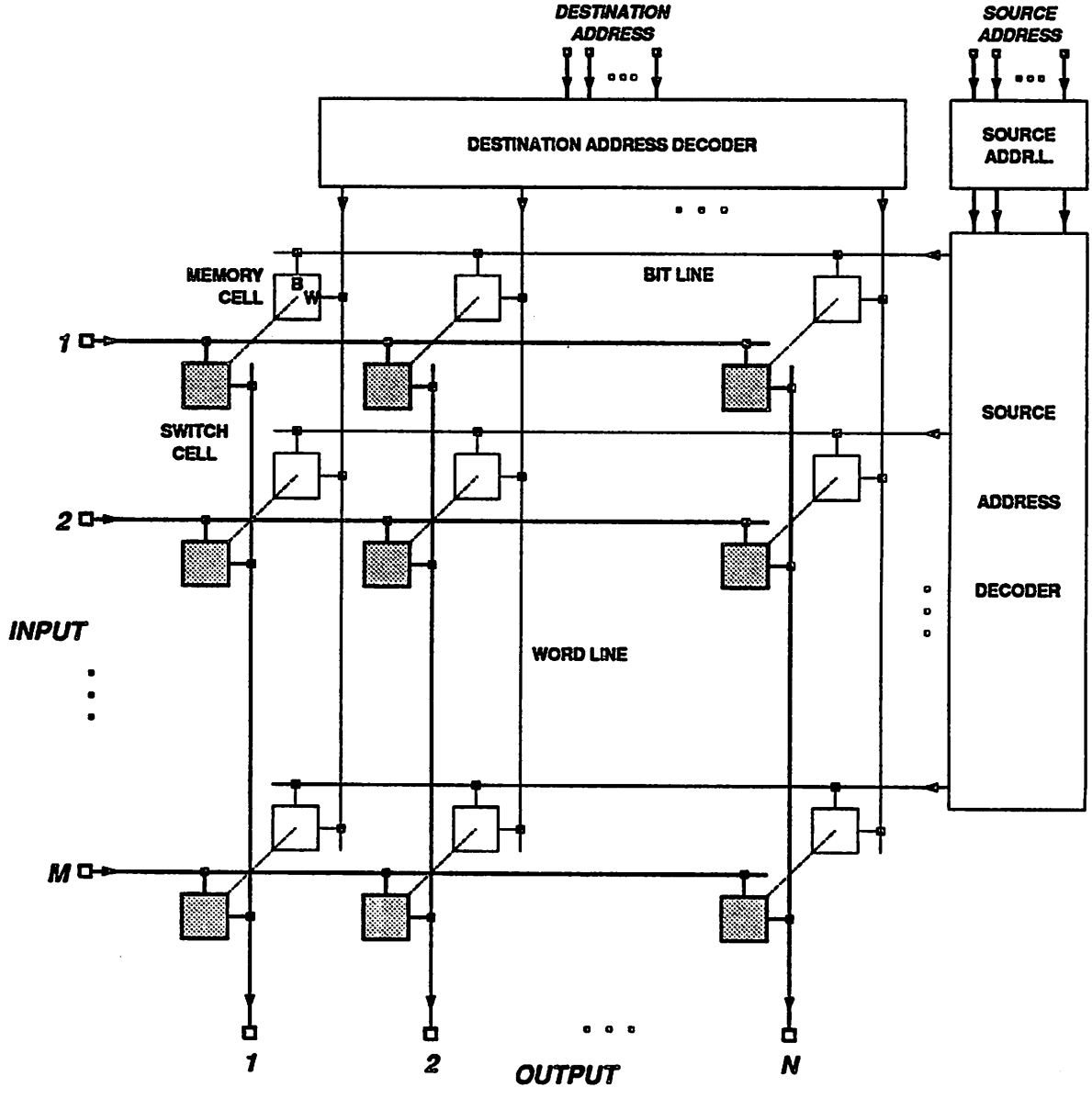


Figure 3.6 Matrix type switch architecture: destinations initiate setup requests.

Figure 3.6 shows the architecture in which switching is initiated by destinations. The decoded destination address line provides gating for all memory cells attached to the corresponding output and the source address decoder generates a bit pattern to be written into these cells. In conventional terms, the destination address line is equivalent to the word line and the source address line corresponds to the bit line.

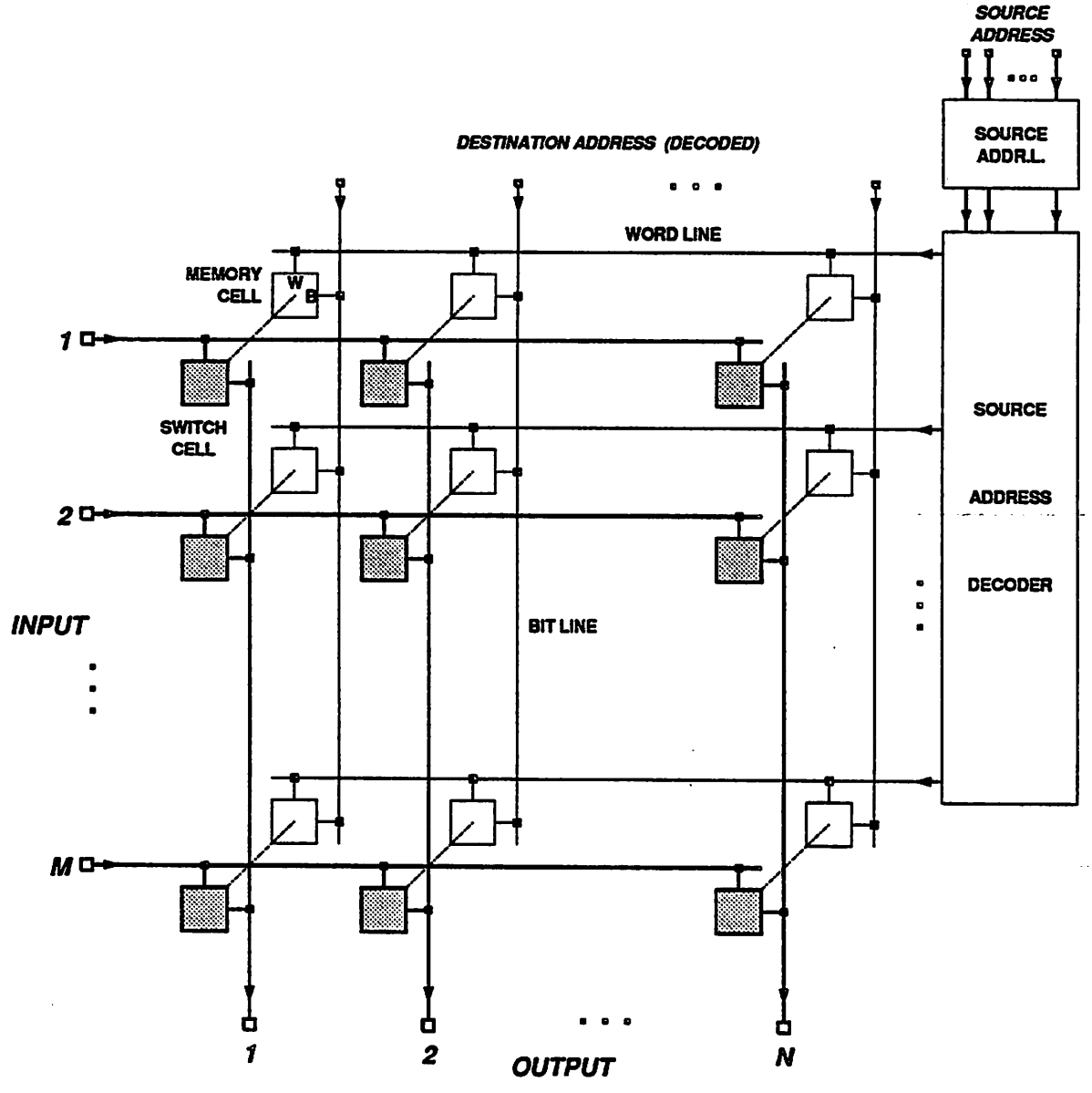


Figure 3.7 Matrix type switch architecture: sources initiate setup requests.

Like the multiplexer type, this architecture naturally prevents two or more sources from occupying one output link and allows at most only one source to have the link. If no source is connected to the link, i.e. the output is idle, all of the memory cells related to the output store '0'. To write this, the source address decoder needs an extra decoding state where its outputs are all low. In addition, this architecture is suitable for implementing switches with bypass inputs as shown in Figure 2.10. Selecting the bypass inputs requires an additional state of decoding. A disadvantage of this destination driven architecture is that it is not convenient for broadcasting. It needs multiple setup cycles to configure for a broadcasting mode.

Another matrix architecture where switching is initiated by sources rather than destinations is depicted in Figure 3.7. In this structure, the source address is decoded to gate all memory cells connected to the corresponding input and the destination addresses are provided on the bit lines. The destination addresses are specified in a fully decoded N bit word that can have as many '1's and '0's. Thus broadcasting configurations require the same number of setup cycles (e.g., one cycle) as that for point-to-point connections. Also the extra source decoding for the idle output state is not necessary. But this architecture needs special hardware or software to prevent two or more memory cells in a column from being written with '1'.

In this thesis, only the matrix type switch architectures are considered for further discussion, because of their advantages over the multiplexer type.

CHAPTER 4

CMOS CROSSPOINT SWITCH

Modern CMOS technologies with channel lengths 2 μm or less meet the bandwidth requirement of a 100 to 200 Mb/s crosspoint switch, as has been shown in Chapter 3. This fact is enlightening because very popular CMOS VLSI technologies, once considered low-speed, can now be applicable for high-speed electronic switching. System integration can be enhanced through VLSI implementation of high-speed switch arrays and control or store functions. However, since the performance margins of CMOS technologies are relatively small, switch design needs careful selection of circuits and techniques to achieve satisfactory performance.

This chapter discusses, first, the selection of CMOS crosspoint cells and memory cells that are important building blocks for matrix type switch architectures. Next, discussions move onto general problems and possible solutions associated with the high-speed, multiple channel switch design.

4.1 CROSSPOINT SWITCH CELLS

Besides satisfying the bandwidth requirement, crosspoint cells in matrix-type switches should have small area and power requirements and must facilitate routing input/outputs and switching signals. In general, switch cells in CMOS design can be broadly classified as passive and active; this distinction only applies when the switches are closed, because the outputs are of high impedance otherwise.

4.1.1 Passive Switches

Passive switches are basically transmission (or pass) gates that provide low resistance direct paths between inputs and outputs when they are ON. As shown in Figure 4.1, either single n-channel transistor or a n-p pair can be used for this type of switches. In these very simple switches, original signal waveforms are maintained throughout the paths except some possible attenuation. Therefore in practice, this type is advantageous for analog signal switching. Especially the fully complementary (n-p pair) switch shows a low resistance over a wide range of signal level and is suitable for large swing applications. However, passive switches cause loading fluctuation on the input nodes regardless of whether they are driven by current or voltage sources. The loading increases as the source is broadcast to more destinations.

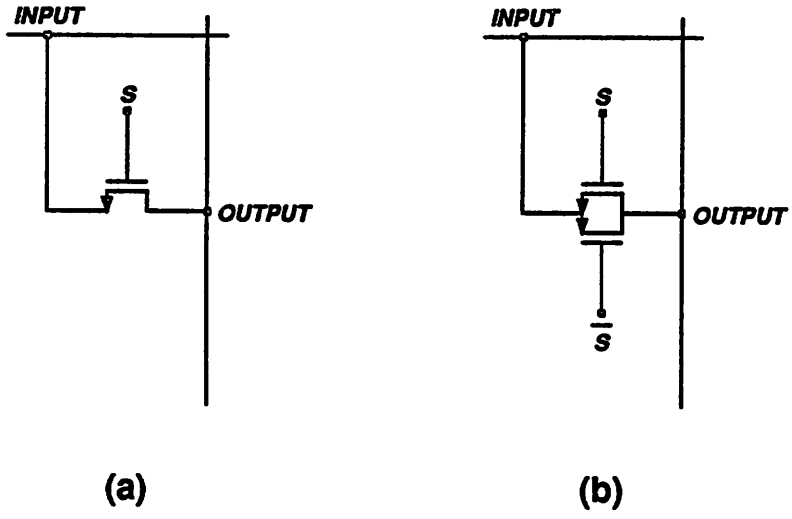


Figure 4.1 Passive switch cells in the form of transmission gate: (a) n-channel MOSFET or (b) n-p pair (fully CMOS) implementation.

4.1.2 Active Switches

Active switches either block or amplify signals. For digital data transmission, the amplification is highly nonlinear and a major reshaping occurs through the switches. The common circuit structures for this type include the standard tri-state CMOS buffer and tri-state CMOS inverters as

depicted in Figures 4.2 and 4.3. In these circuits, the tri-state control terminal is connected to a memory that contains the switching information.

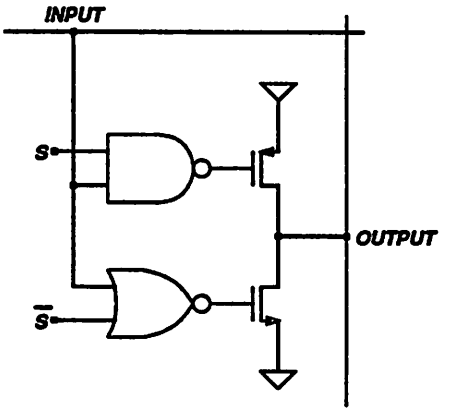


Figure 4.2 Active switch cell: standard tri-state CMOS buffer implementation.

The standard buffer (Figure 4.2) does not have stacked devices in the output circuitry, so it is good for driving a heavy load. However, the logic gates in the signal paths add to the delay through this switch. Furthermore, if the delays in the NAND and NOR gates differ, skew will be added to the signal. Another thing to note is that the total number of devices is 10. Hence, this switch cell would consume much area and power.

The tri-state inverters in Figure 4.3 have only 4 devices each while achieving the same function and solve many of the problems associated with the standard buffer. However, due to stacking of devices, to maintain a drive capacity equal to that of the standard buffer, device sizes must be increased by about a factor of 2. Although all three circuits are functionally equivalent, their minor differences are significant for high-speed data switching.

The circuit (a) is different from the others in that the switching transistors (MSN, MSP) are connected to the power supplies and the inverting transistors (MIN, MIP) are attached to the output node. As a result, the input signal is coupled to the output node through the gate-drain overlap capacitances of MIN and MIP, even though the switch is OFF. This leads to severe crosstalk

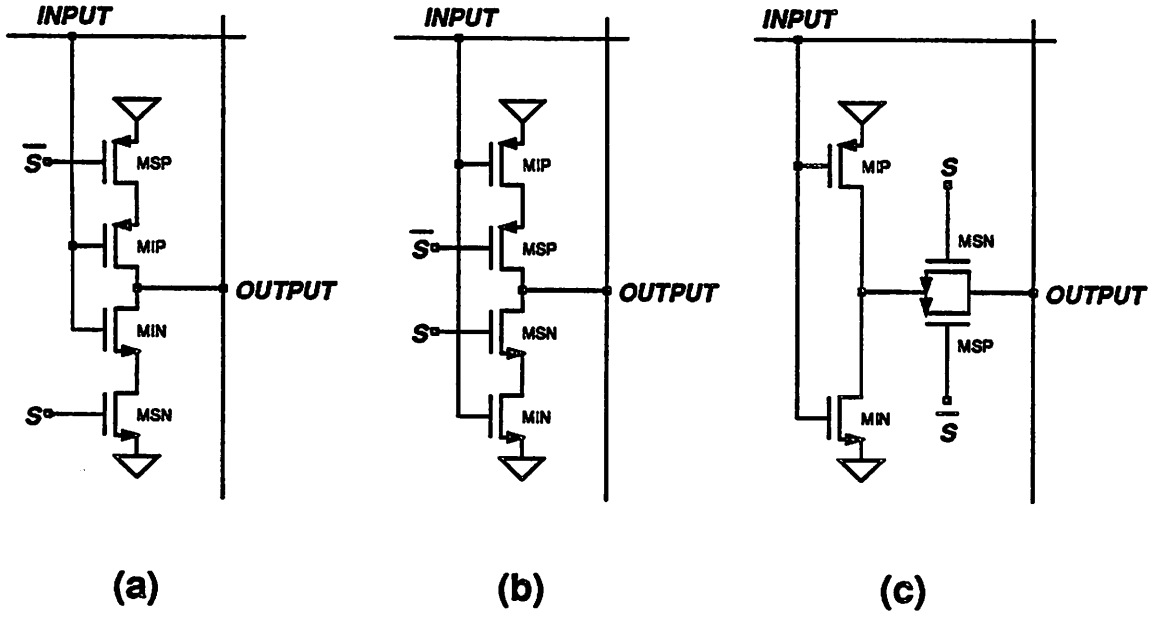


Figure 4.3 Active switch cells in the form of tri-state CMOS inverter: switching transistors (MSN and MSP) are (a) not connected, (b) connected, and (c) connected as a transmission gate to the output.

on one active output node from the switch cells that are connected to it but turned OFF. Also, because one of the inverting transistors is always ON, a switch that is ON should charge the nodes inside other switches that are OFF, which the switching and inverting transistors share (i.e., drain of MSN or drain of MSP). Net effects of this are the increase of output loading and decrease in switch bandwidth.

The circuits (b) and (c) are advantageous in these respects. Because the switching transistors are located between the output node and inverting transistors, if the switch is OFF, the feed-through paths from the input to output are effectively blocked. Therefore, both crosstalk and unwanted loading on the output are minimized. The choice between (b) and (c) depends on the applications. Generally, (c) has more driving capability because the n-p pair transmission gate shows less ON resistance if the same size devices are used. However, it requires more area due to additional contacts and consumes power in the inverter even when the switch is OFF. A more

compact layout results for (b) and power is not dissipated in the switches that are OFF.

4.2 CMOS MEMORY CELLS

Memory cells for the matrix type switch need to be compact and should not complicate the overall routing. In this sense, the choice of memory cells is based on the number of devices and signal lines. Actually, more weight should be put on minimizing the number of lines since there are already input, output, and two power supply lines to be routed in a crosspoint/memory combined cell.

4.2.1 Standard Static RAM and Latches

One common memory cell is the CMOS static *random-access-memory* (RAM) cell as shown in Figure 4.4 (a), which has only 6 transistors, one word line and two bit lines. Its area efficiency makes it popular for most static RAM implementation. But this cell would not be appropriate for matrix type switch application due to routing difficulty. For proper writing, the transistor sizes must be carefully determined because signals on the bit lines after passing the transfer gates should override the opposite forces in the memory.

Another common circuit that can be used for memory cell is the CMOS standard latch (Figure 4.4 (b)). This latch consists of 8 transistors, one bit line, and two word lines. In terms of device and line counts, this cell has two more transistors than the static RAM cell. Consequently, this type needs more area. An advantage of this cell is that the device sizing is not critical. This is because the feedback path is open when writing the cell; no node is driven by two low impedance sources at the same time. However, a severe transient skew between the two complementary signals on the word lines can close the feedback path at the time of writing and cause conflict. Thus the transient skew must be minimized.

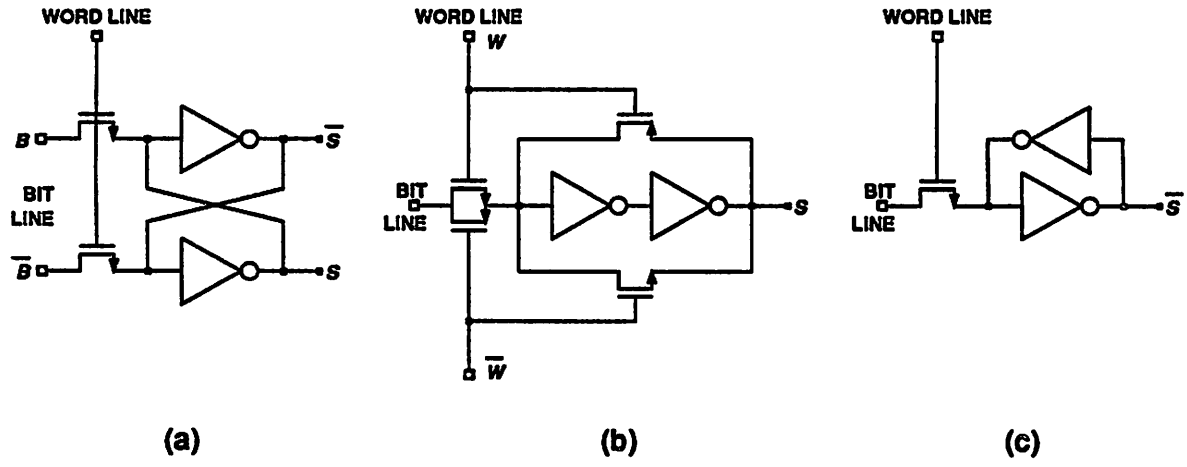


Figure 4.4 Memory cells: (a) CMOS static RAM cell, (b) standard static latch, and (c) a variation of static latch.

A variation of the standard latch shown in Figure 4.4 (c) is a good choice because of its small number of devices and lines: 5 transistors, one word line, and one bit line. It can also be thought of as a modification of the RAM cell with one transfer gate removed. This cell is most advantageous in minimizing area and routing complexity. But like the RAM cell, aspect ratios of devices need to be adjusted for proper operation. Normally, the inverter in the feedback path uses long and narrow transistors that are larger than minimum size ones. This predicts the cell area larger than expected from the device counts.

4.2.2 A Modified Static Latch

Another form of latch depicted in Figure 4.5 can be used as the memory cell. This latch results from removing two transistors connected to the complementary word line from the standard latch. There are 6 transistors, 1 bit line, and 1 word line in this cell. It occupies a small area and is easily routed. Other benefits of this cell include non-critical circuit sizing and freedom from the skew problem. Additionally, minimum size transistors can be used to reduce area. However, this latch has less noise immunity and has some static power consumption. A short pulse width for the word line may help reduce the waste of power.

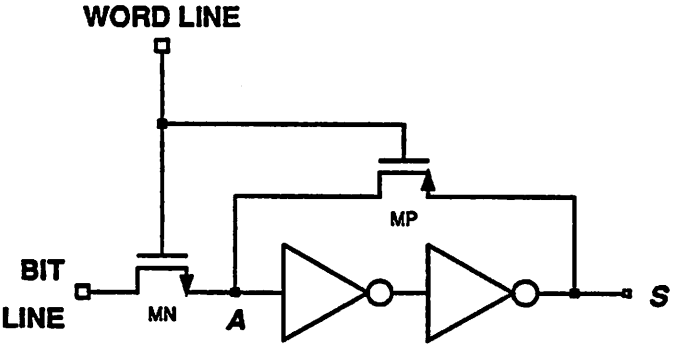


Figure 4.5 A modified static latch memory cell.

When writing '1' into the cell storing '0', the node *A* gets charged to $V_{DD} - V_{Tn}$, where V_{Tn} is the threshold voltage of MN with the body effect. If this voltage is greater than the logic threshold voltage of the first inverter, the logic level propagates through the inverters adequately. But if the voltage is lower than $V_{DD} - |V_{T0p}|$, where V_{T0p} is the zero bias threshold voltage of p-MOSFETs, the p-MOSFET in the first inverter conducts and a current flows directly from V_{DD} to V_{SS}, although the first inverter output remains low. This lasts as long as the word line is high. As soon as the line goes low, the feedback transistor MP becomes ON and pulls *A* to V_{DD}. Therefore, the power consumption stops afterwards. As a result, the narrower the gating pulse on the word line, the less the power dissipation.

The above problem does not happen when writing '0', since the node *A* is fully discharged to V_{SS}. However in latching mode where the word line is low, a different kind of problem occurs. Because the source and drain of MP are at ground potential, MP is effectively turned OFF. In this case, *A* becomes a floating node that is vulnerable to any noise coupling. A noise voltage up to $|V_{Tp}|$ (the threshold voltage of MP with the body effect) can build up on this node. Beyond that, MP turns ON and extra charge is absorbed. Therefore, the logic threshold voltage of the first inverter must stay between $V_{DD} - V_{Tn}$ and $V_{SS} + |V_{Tp}|$. Provided that $|V_{Tp}| > V_{T0n}$, there is a possibility that a small current flows from V_{DD} to V_{SS} through the first inverter. V_{T0n} is the

threshold voltage of n-MOSFETs without a body bias.

This potential noise and power dissipation can be minimized by reducing the impedance of node *A* and amount of noise coupling. One trick would be to use leaky p-MOSFETs that do not completely turn OFF. P-MOSFETs in ordinary CMOS technologies may have this leaky characteristic due to the parasitic buried channel.

4.3 CMOS SWITCH DESIGN PROBLEMS

Major problems in designing electronic switches are associated with high-speed signal paths. At high speeds, effects of the large voltage swing and large transient current in combination with capacitive or inductive parasitics in the signal paths, which are unnoticeable at low-speeds, become apparent. Data propagating through the matrix of crosspoints may experience interferences in the forms of crosstalk, intersymbol interference, noise, etc. The degree of influence depends on the bit rate, voltage swing, and transient current as well as on the switching information.

4.3.1 Large Voltage Swing and Transient Current

First thing to note is the bit rate that goes up to 200 Mb/s or higher. To achieve such a high bit rate, as mentioned in Section 3.3.3, the slew rate dv/dt for each output transient must be large. This is mainly because Δt gets small for a unit voltage swing. Secondly, the internal voltage swing for conventional digital CMOS ICs is quite large, typically rail-to-rail (5 V). Because of this large swing, noise coupled onto other circuits through capacitances will be large.

Due to this high slew rate requirement, in CMOS circuits where loads are primarily capacitive, driving or charging currents

$$i = C \frac{dv}{dt} \tag{4.1}$$

should be large during transients. A large driving current during a transient necessarily implies that the rate of change di/dt is also large. Therefore, any parasitic inductances along the current paths induce substantial voltages

$$v = L \frac{di}{dt} \tag{4.2}$$

across them.

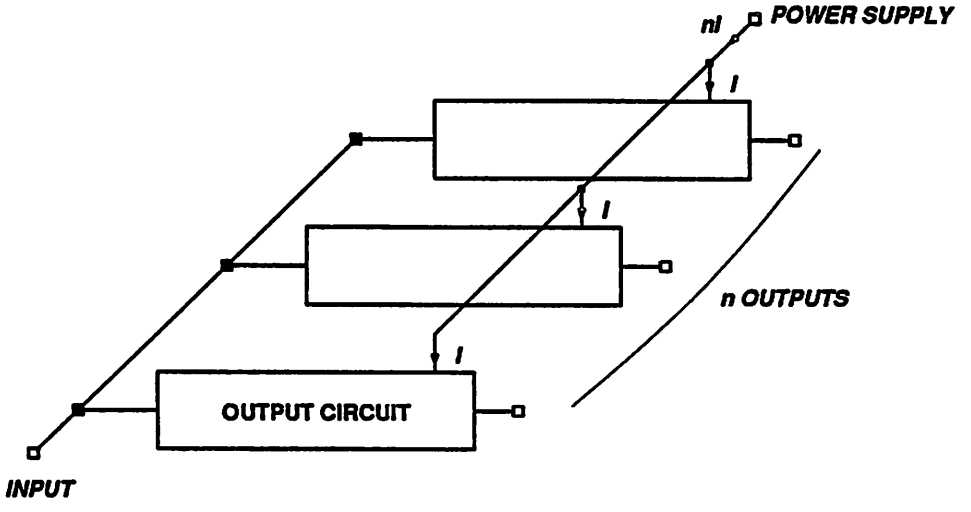


Figure 4.6 When one input is broadcast to n outputs, the total transient current from the power supply increases n-fold.

Third thing to note is that many outputs change simultaneously when broadcasting, where one source station is connected to multiple destinations up to N. If one input is broadcast to n outputs, the input signal is switched to and propagates through n identical output circuits at the same time. Because n circuits have simultaneous voltage and current transitions of equal amount, the total transient current from the power supply is n-times the transient current in one output circuit as shown in Figure 4.6. Hence, the voltage drop on any parasitic inductance along the power supply line is n-times larger than the drop for one active output. The drop becomes worse as n approaches N and is dependent upon the switching information.

4.3.2 Capacitive and Inductive Parasitics

The amount of crosstalk and jitter in the high-speed paths of a switch is directly proportional to the capacitive and inductive parasitics. Parasitic capacitances are abundant in the matrix area of the switch. As illustrated in Figure 4.7, the matrix area consists of M input and N output lines that extend the entire horizontal and vertical dimensions of the matrix, respectively, and array of crosspoint cells. Since each input line is connected to N crosspoint cells, if one of the CMOS cells shown in Figures 4.2 and 4.3 is used, it imposes a large capacitance onto the previous circuit such as an input buffer. This loading capacitance comes from the gate, gate-drain overlap with the Miller effect, and interconnection. Similarly, every output line connected to M cells has a loading capacitance originating from drain junction, drain-gate overlap, interconnection, and the next stage like an output buffer. The loading capacitances degrade the bandwidth of the high-speed paths and thereby increase the intersymbol interference or pattern-dependent timing jitter. Also, there are coupling capacitances among the input lines and among the output lines, which cause crosstalk between channels. The input-output overlap (and gate-drain overlap in some cells) presents a coupling capacitance between input and output lines and causes crosstalk via feedthrough when the cell is OFF.

The matrix area also has inductive parasitics, especially on the long input/output lines. However, the inductance values for typical, on-chip metal routing are relatively small. For example, 1 μm thick, 2 μm wide, and 2 mm long aluminum line has an inductance around 3 nH (Figure 4.8) [41]. Assuming that the current transients on the input and output lines are as high as 10 mA/ns, the voltage drop on this line is about 30 mV. Hence, the effects of these inductances are unnoticeable compared to the capacitance or power supply inductance effects.

The parasitics related to the power supply nodes on the chip have more pronounced effects than those in the matrix in terms of interchannel crosstalk and intersymbol interference. Figure 4.9 depicts these parasitics from input section to output section of the chip referenced to the

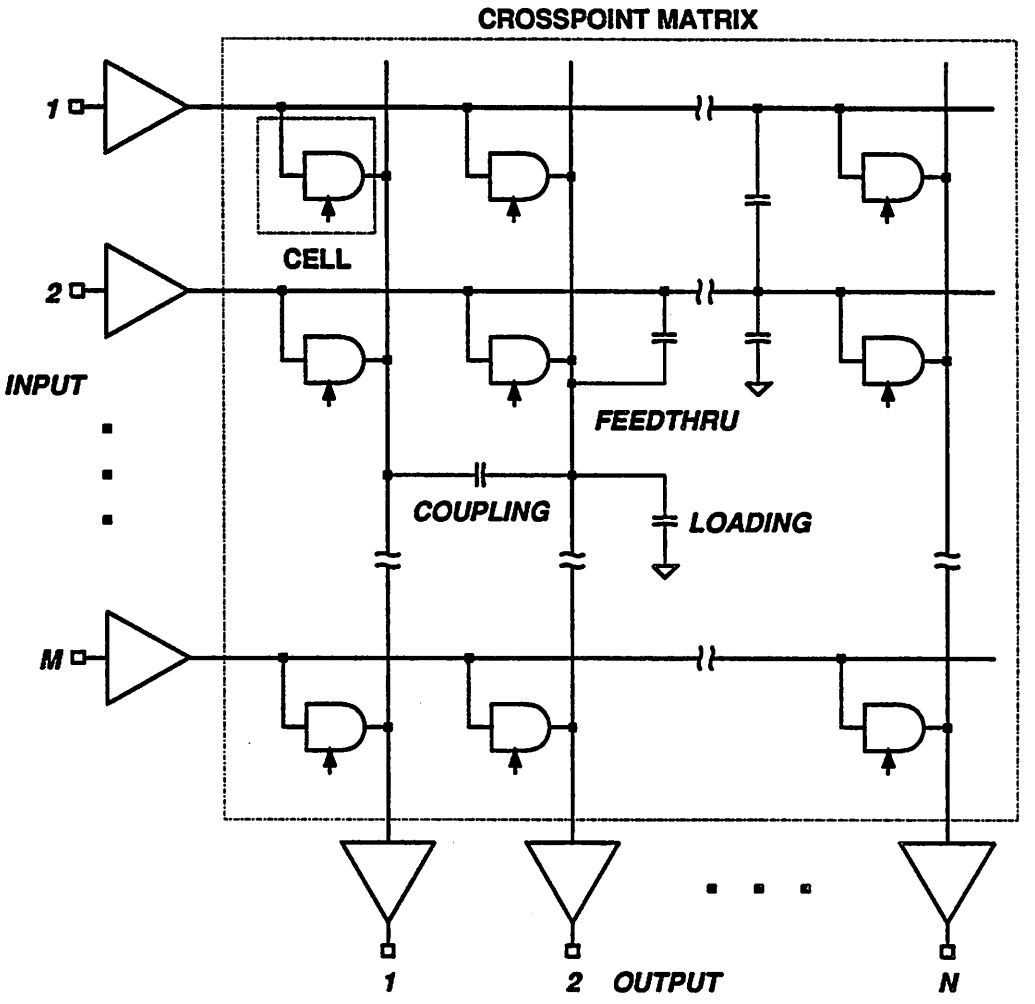


Figure 4.7 Capacitive parasitics in the matrix area.

external nodes such as power supply ([V_{DD}]) and ground ([V_{SS}]). Because a chip is connected to the external world mostly through bonding wires (between on-chip pads and package bonding pads) and package leads, every input, output, and power supply line has parasitic inductance associated with the packaging. Typical inductance values for various package leads range from 5 to 20 nH and the inductance of bonding wires is about 1 nH [42,43].

As has been mentioned, when the switch is in a broadcasting mode, many outputs move synchronously and transient currents add up (Figure 4.6) to flow through the power supplies. The largest part of the transient currents may come from the output buffers that drive external loads.

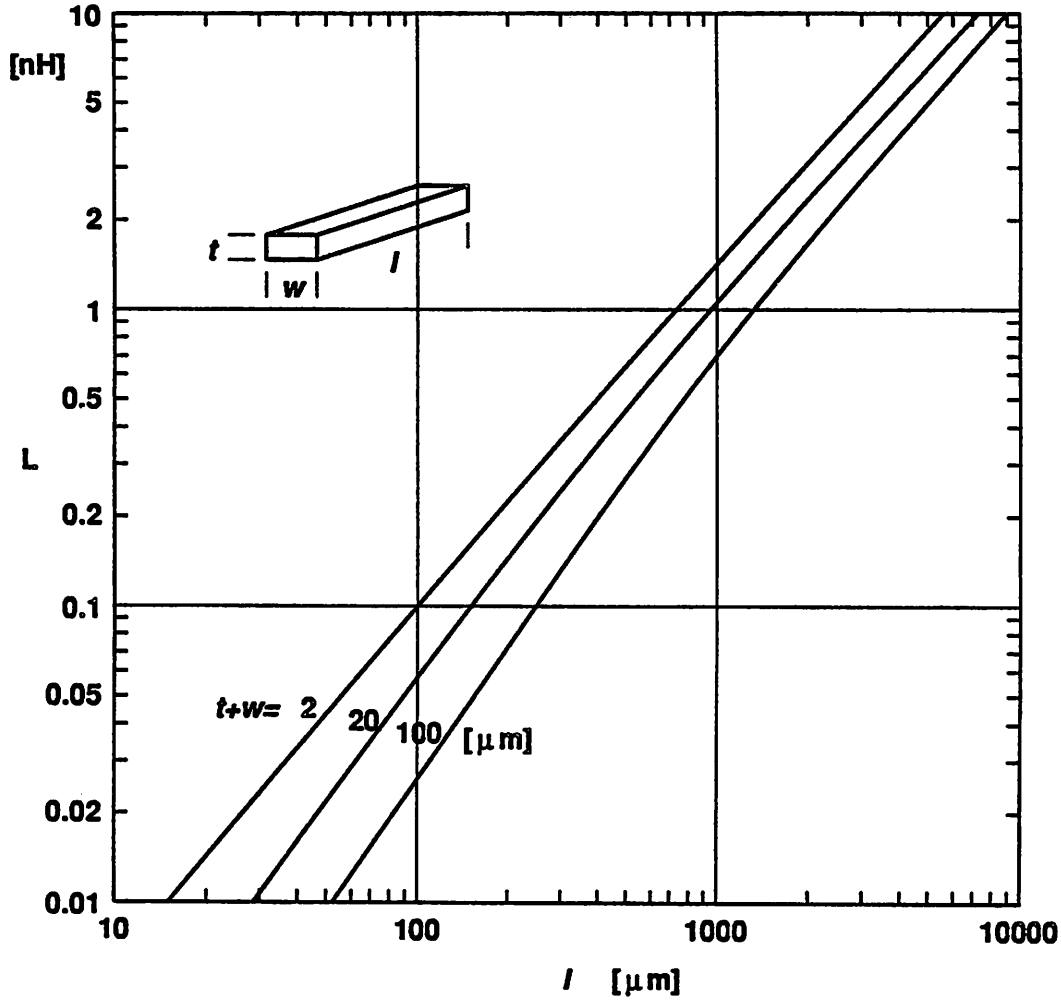


Figure 4.8 Self inductance of rectangular straight wire.

The voltage drops on the power supply inductances increase proportionate to the number of destinations. The worst case when N outputs have simultaneous transitions with a common power supply line can be equivalently considered as the case when one output makes transitions with an effective power supply inductance N times larger than the real value. Therefore, inductances in the power supply leads are more harmful than those in the input and output leads, if their values are comparable.

Another thing to note in Figure 4.9 is the capacitances from internal nodes to the on-chip power supplies and capacitance between the supplies. Due to these capacitances, any voltage

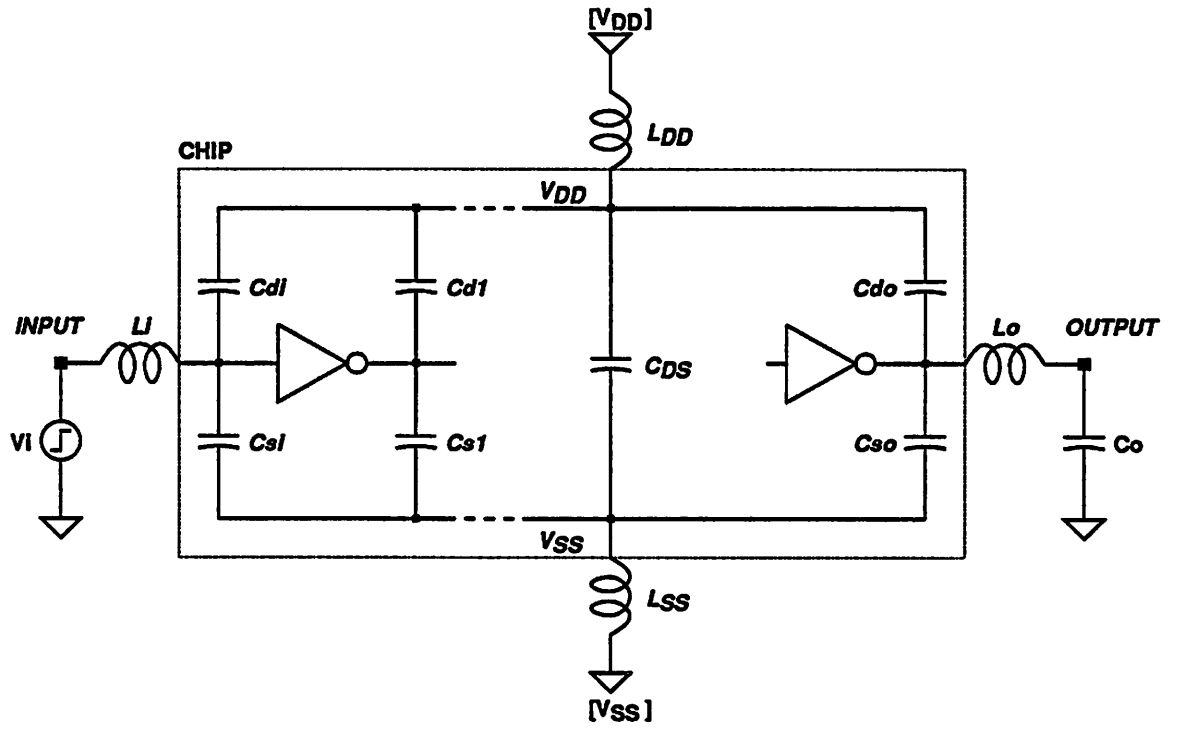


Figure 4.9 Capacitive and inductive parasitics related to the on-chip power supplies, inputs, and outputs.

swing on the nodes is coupled to the non-perfect on-chip power supply nodes and vise versa. Especially, a swing on the input node directly moves the on-chip power supply nodes (V_{DD} and V_{SS}) up and down.

The current transients acting on the power supply inductances as well as the capacitive coupling from the internal nodes cause on-chip supply fluctuation. Then, since the internal nodes are connected to the on-chip supplies through capacitances or devices that are ON, they are influenced by the supply voltage fluctuation. Because this is a RLC circuit in essence, certain nodes in the circuit ring. The ringing amplitude decays with time. The amplitude, frequency, and decay time of ringing are determined by all capacitances, inductances, and resistances of devices that are ON. In general, the resistances and capacitances vary instantaneously with the supply and internal node fluctuation.

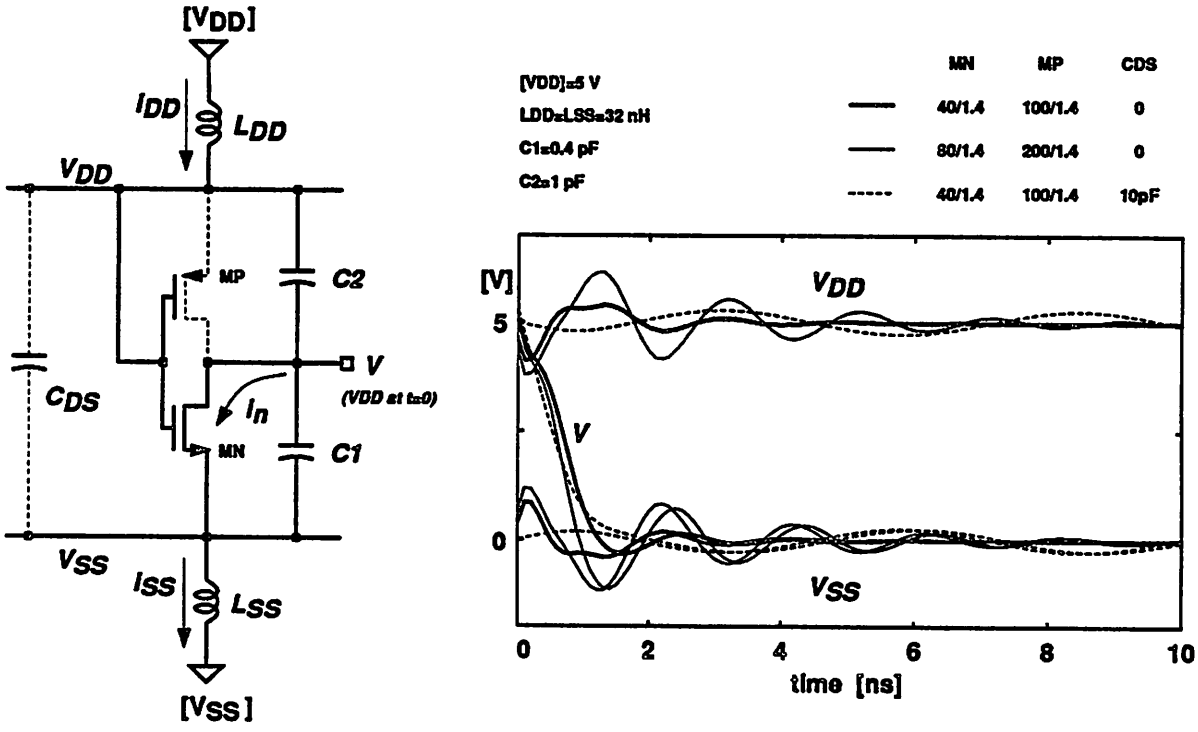


Figure 4.10 Simulation results on the effects of parasitic inductances and capacitances.

As an example, let us consider a simple inverter stage that is discharging output node from V_{DD} (Figure 4.10). The output node has two parasitic capacitors connected to the on-chip power supplies. As soon as the input of the inverter is tied to V_{DD}, the n-channel transistor is ON and current changes rapidly. Between the two components of current flowing through the transistor, only the one from the top capacitor C_2 flows through the supply and induce voltages across the two inductors. Ignoring the effects of inductances and the voltage dependency of capacitances for a moment, the currents flowing through L_{DD} and L_{SS} are

$$i_{DD} = C_2 \frac{d(V_{DD} - v)}{dt} = -C_2 \frac{dv}{dt} \tag{4.3}$$

and

$$i_{SS} = i_n + C_1 \frac{dv}{dt}, \tag{4.4}$$

respectively, where i_n represents the n-MOSFET current. Since $i_{DD} = i_{SS}$, from the two equations,

$$i_{DD} = i_{SS} = \frac{C_2}{C_1 + C_2} i_n. \tag{4.5}$$

Therefore, the supply current during transient is directly proportional to the transistor current that is varying following the I_{DS} vs. V_{DS} curve [44]. This proportionality is retained even when the input is connected to V_{SS}; that is,

$$i_{DD} = i_{SS} = \frac{C_1}{C_1 + C_2} i_p. \tag{4.6}$$

where i_p is the p-MOSFET current. The relationships become more complicated in real CMOS circuits, because the capacitances are voltage dependent and, hence, the proportionality factor is a function of voltage or, equivalently, time.

The direction of current flow suggests an equal amount of lift on V_{SS} and dip on V_{DD} for a transient. Thus initially, internal power supply voltage decreases and current drive from the transistor drops; the transistor may be still in the saturation region and drain current depends on $(V_{DD} - V_{SS} - V_{Tn})^2$. Because of this current reduction, discharging slows down and the polarity of the induced voltages reverses. Also, LC ringing triggered by the initial current transient continues. Once the supply voltage increases, the drive current becomes large again (or resistance gets small if the transistor is in the linear region) and the above process repeats until the circuit settles. Settling time for the circuit including power supplies may be quite long. So, if the input connection changes before settling, new output transition will be different from the previous one. Obviously, a larger device increases the ringing amplitude and settling time. If a bypass capacitance between V_{DD} and V_{SS} is added, the amplitude and frequency of oscillation decrease.

The effects of the complicated interaction discussed above appear as interchannel crosstalk and intersymbol interference. Or equivalently, it degrades data with pattern-dependent timing jitter and noise. In analog circuit terminology, the circuit may be said to have a poor power supply rejection.

In addition, the fluctuation of power supply nodes inside the chip and severe noise on the internal nodes may cause latch-up in CMOS. P-n junctions may be forward biased and inject carriers that can trigger latch-up.

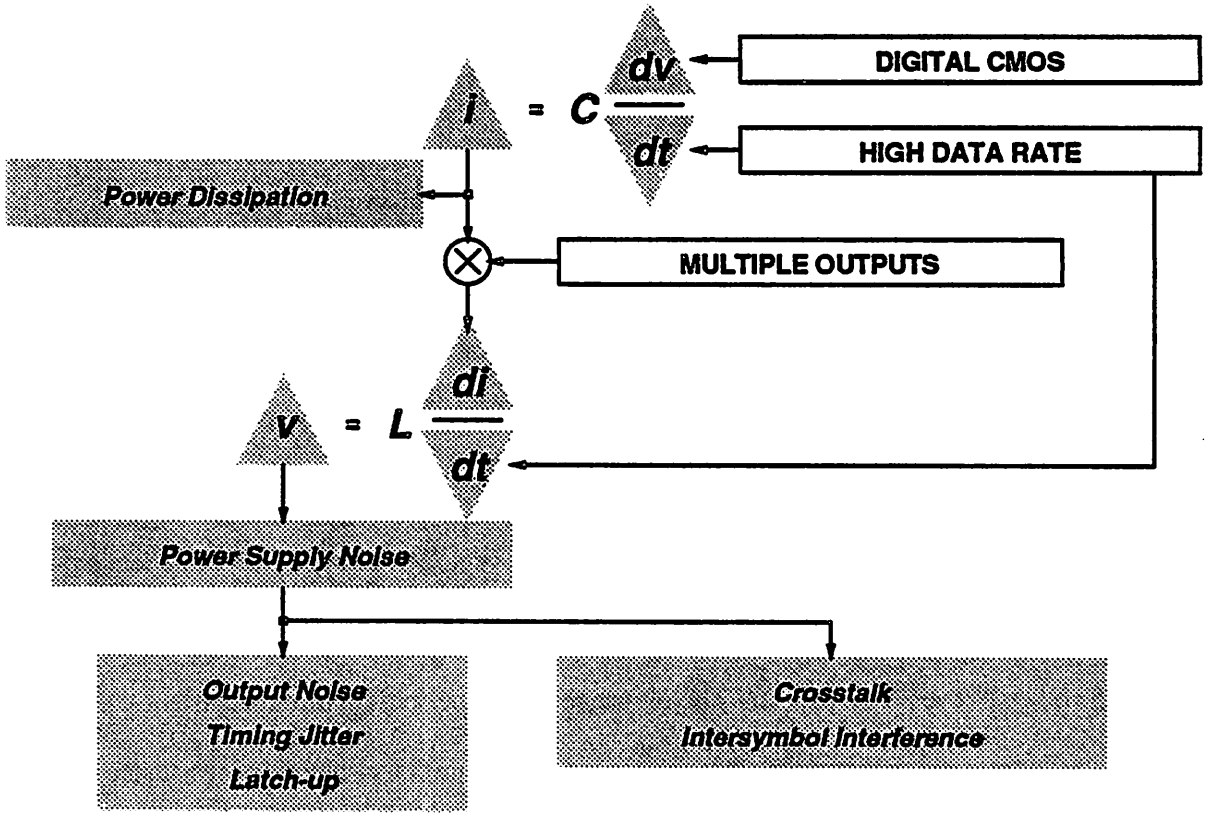


Figure 4.11 Summary of high-speed CMOS switch design problems.

Aside from the output timing jitter and noise, and power supply noise caused by the parasitics, a high-speed switch consumes large power either static or dynamic. If the switch is implemented as a conventional digital CMOS circuit, the power consumption is mostly dynamic; dynamic power is represented by $CV_{DD}^2 f$. Because output loading capacitance is generally much larger than on-chip capacitances, most of the dynamic power is dissipated at the output stage. For example, a 5 V, 16 × 16, 200 Mb/s switch with total 20 pF capacitance along each high-speed path consumes 50 mW of dynamic power, on one path. When all outputs are active, the power

sums up to 800 mW. If the number of channels to be switched increases, more expensive packaging may be needed for adequate chip cooling.

Figure 4.11 summarizes the problems in designing a high-speed CMOS switch. The significance of the problems discussed in this section is related to ensuring error-free propagation of data at the required speed, not merely achieving the speed. If the technology has enough bandwidth, CMOS switches designed in standard digital techniques can pass high-speed data. But the data waveforms coming out of the switches will be badly damaged with noise and timing jitter that may cause data errors. Since the noise and jitter levels must be limited for proper data transmission, design techniques to reduce these levels are needed.

4.3.3 Input and Output Interface

Proper interface between the input and output of the switch is necessary to enable switch network expansion. In CMOS, this interface is likely to be digital CMOS one that is specified as voltage levels close to V_{DD} and V_{SS}. But, as mentioned before, high input/output swings are bad because of large noise caused by them. Especially, if the output stage is not fully differential, the supply current fluctuates heavily during transients and unbalance increases between i_{DD} and i_{SS}. It would be better to have low-swing voltage interface or even current interface, to reduce the noise level. To minimize the effects of input and output inductances, however, the current swing in a current interface should also be small.

Making the input and output compatible with external standard logic elements, opto-electric receivers, and electro-optic transmitters is another design consideration. Because bipolar *emitter-coupled logic* (ECL) is designed for high-speed applications (e.g., 100 Mb/s or higher), most fiber-optic receivers and transmitters have been designed to be ECL compatible [28]. Therefore, if input/output levels of the switch are compatible with the ECL levels, the switch would be easily interfaced to other components in a system. In addition, since the ECL swing is

only about 1 V, the noise coupled from the input node to the circuits inside will be small compared to, say, 5 V interfaces. The noise induced from driving the output would be also less.

However, it is not easy to implement ECL output levels in CMOS technologies, due to the lack of good bipolar transistors and resistors. Only approximated levels that are not precisely controllable and have different temperature characteristics can be obtained using CMOS circuits. ECL compatible inputs can be designed with less difficulty using differential circuits.

4.4 CMOS SWITCH DESIGN TECHNIQUES

There are several design techniques that improve the noise and jitter levels in a switch while maintaining the desired speed. One obvious solution to these problems is to reduce the parasitics directly. Other solutions would involve circuit techniques which cut down device sizes, transient currents, and voltage swings. Because each technique may have its own limitation, to get a maximum performance, it would be better to combine several compatible techniques.

4.4.1 Reduction of Capacitive Parasitics

For a given technology, on-chip parasitic capacitances can be directly minimized by using small geometry devices, small drain junction area, and minimum width wires and by optimizing the layout so that interconnection lengths and overlap area is minimized.

Sizes of the devices may not be reduced freely because the speed requirement of the switch puts constraints on the device aspect ratios. The best choice will be to use minimum channel length devices wherever possible. Then, gate capacitance and gate-drain overlap capacitance are minimized. But in some cases, channel lengths cannot be decreased further because either device aspect ratios less than 1 (minimum width must be used instead of minimum length) or large drain-source resistances are required.

Drain junction area that normally includes contact area can be reduced by sharing the area with other devices, using minimum diffusion widths, and minimizing contact area, as shown in Figure 4.12. For one big device, it is better to divide it into many small devices of straight or serpentine form, place in parallel, and share the drain area. However, use of the minimum diffusion widths increases the parasitic resistances from the active drain edges to the contacts. Also, contact area may not be reduced further because of the current density limit of contact or the contact resistance. A big device needs large contact area or a large number of unit contacts to deliver the current efficiently. Typically the maximum current per minimum-size unit contact is about 1 mA.

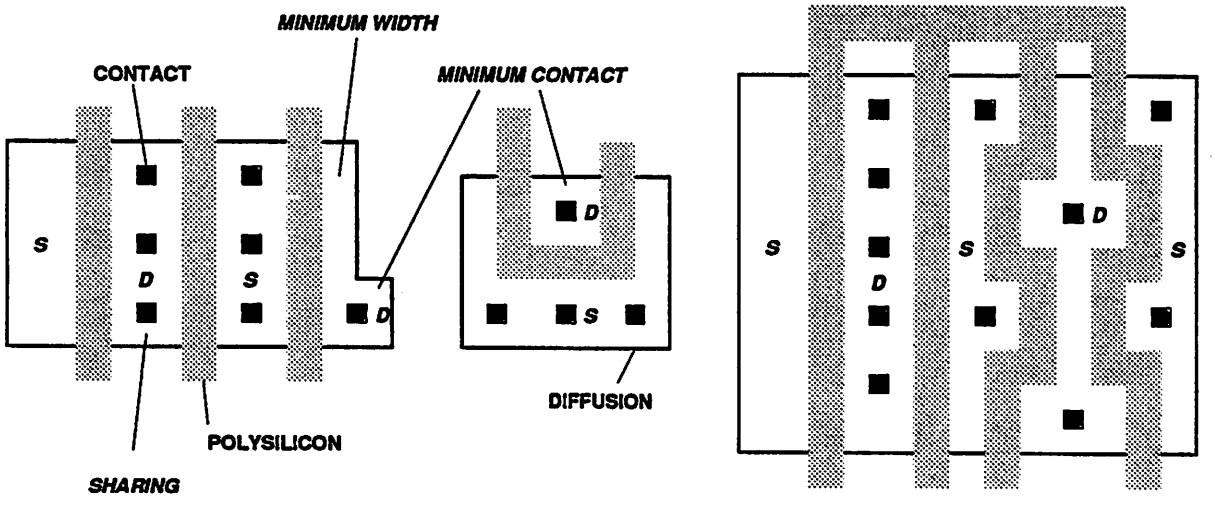


Figure 4.12 Layout examples with reduced drain area.

Interconnection capacitance for a fixed length is proportional to the wire width and inversely proportional to the dielectric material thickness, if fringe capacitance is ignored. Thus, the capacitance decreases when a top level layer (e.g., second metal in double-metal processes) with a minimum width is used for long signal routing. But due to increased proportion of the fringe capacitance as the line width is reduced, the benefit becomes less noticeable. Another thing that limits minimum line width is the maximum current density of the wire before severe

electromigration happens; typical maximum current density for aluminum layer is around 1 mA per 1 μm width. If metal layers are used for routing, parasitic resistance along the line is negligible. In addition, since rate of increase of inductance is much lower than rate of decrease of capacitance when the line width is reduced, it is advantageous to use narrower wires (see Figure 4.8).

4.4.2 Reduction of Inductive Parasitics

As mentioned in Section 4.3.2, on-chip inductances are negligible compared to the inductances associated with packaging. Because transient current level in internal branches is usually lower than that in final output circuits, effect of the on-chip inductances would be relatively unnoticeable. If on-chip inductances are still large, they can be reduced by widening the wire widths. However, rapid increase in capacitances would do more harm than good. Since the capacitance increase is advantageous for power rails, only internal power supply routing will benefit from this widening.

The unwanted inductances from packaging will be reduced by using short-lead packages, multiple bonding, and multiple pads. Although the values vary depending on the number of pins, cavity size, and particular design, short-lead packages such as *leadless chip-carriers* (LCCs) and *pin-grid arrays* (PGAs) have smaller inductances than most common *dual-in-line packages* (DIPs). For example, a 50 mil pin-spacing, 64-pin LCC shows inductances ranging from 8 to 10 nH [42] and a 64-pin PGA has inductance distribution of 6 to 13 nH [43]. Generally, LCCs show smaller variation of inductances from pin to pin and lower average values than PGAs, especially if pin count is less than 64. Also, flip-chip packaging technology that uses solder balls (or bumps) for the flip-chip interconnect ensures very low inductances [45].

To reduce packaging inductance as a whole for a certain signal path, multiple pads both on the chip and package can be assigned to the particular node and multiple bonding can be done, as depicted in Figure 4.13. This multiple bonding will cut the packaging inductance in proportion to

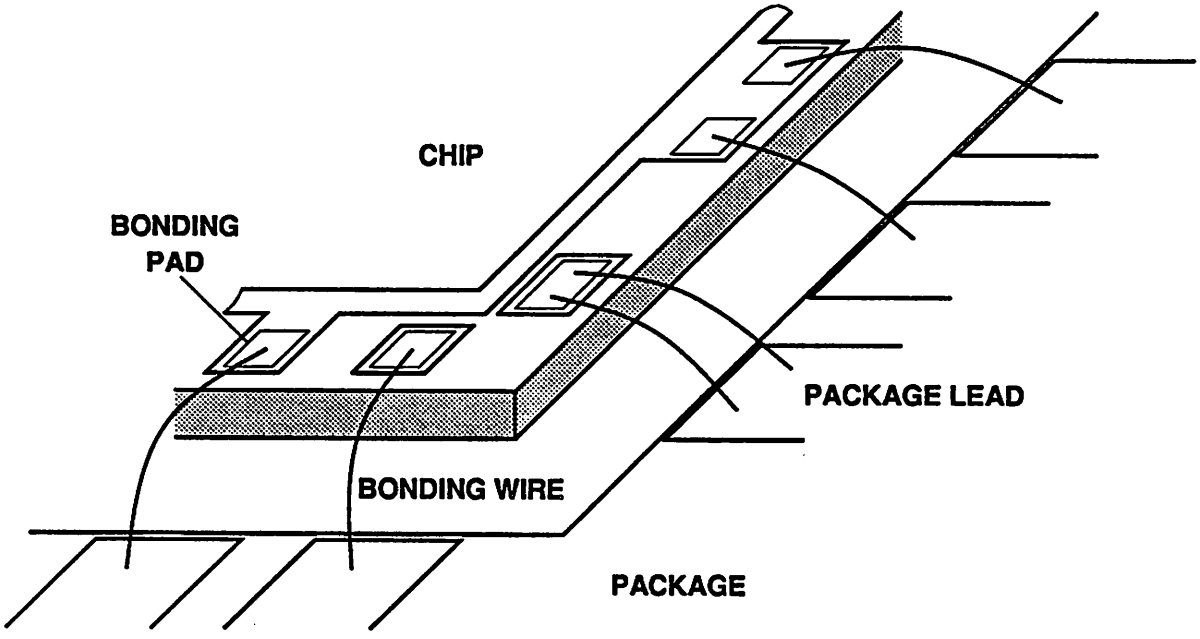


Figure 4.13 Multiple bonding to reduce packaging inductances.

the number of bonding wires. In general, power supply nodes require this technique, since the effect of the inductances is severe. However, assigning multiple pads is followed by the increase in the chip area, if the area is determined by the pad pitch and number of pads. Then, it inevitably adds on-chip capacitances and inductances. As well, it implies the use of more expensive packages with larger numbers of pins and longer leads. Another form of multiple bonding between one pad on the chip and one pad on the package can also be used to reduce the bonding inductance. Even though lengths of the bonding wires are not easily controllable because they are determined by the cavity size of packages and the chip size, short bonding will reduce the bonding inductances.

4.4.3 On-Chip Bypass Capacitor

As has been indicated in Figure 4.10, the ringing or noise on the on-chip power supplies due to the parasitic inductance and capacitance can be reduced by using bypass capacitors between the

two on-chip supply nodes. This is because bypass capacitors provide a part of the charging current that flows through the load capacitor and transistor during each transient and, thereby, reduce the current from the external supplies through the parasitic inductors (Figure 4.14).

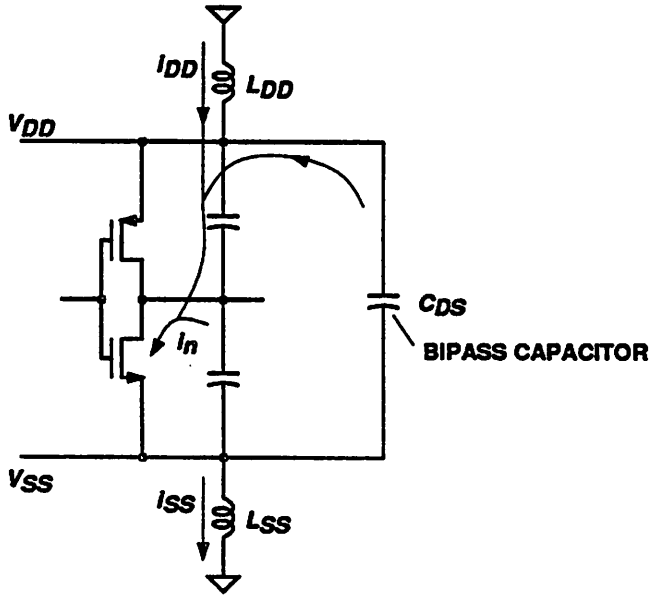


Figure 4.14 On-chip bypass capacitor reducing power supply inductance effects.

In CMOS, reverse-biased, well-to-substrate junction capacitance and overlap capacitance between the power routing layers naturally provide on-chip bypass capacitance. Hence, an effective layout technique for adding bypass capacitance is to increase the well area or cover more area with power supply rails. To be effective, however, the substrate and well contact need to be solid so that series resistance from the power supply to the actual capacitor plate is small.

If the technology provides extra layers such as second poly-silicon, second metal, or third metal, they can be utilized for bypass capacitors. These capacitors formed in any unused area on the chip will not increase unwanted capacitance to nor interfere with other nodes. It may be possible to cover the whole chip with one of these layers and form a large capacitor to substrate. The substrate is connected to V_{DD} in p-well CMOS technologies. But this method will inevitably

increase the capacitance between the internal nodes and power supply nodes.

Bypass capacitors can be connected between the power supply nodes external to the chip but inside the package. Discrete capacitors of substantial capacitance may be placed between the two ends of the package leads where bonding wires are attached. For the bypass capacitance sufficiently large, net parasitic inductance seen from the on-chip supplies is only that of the bonding wires and reduced by as much as that of the package leads.

Because typical bonding wire inductance is about 1 nH and package lead inductance varies from 5 to 20 nH, this on-package bypassing is very effective in reducing the power supply inductance problems. However, it is quite inconvenient to solder these capacitors and it may require special packages. Also, since the capacitors have their own internal lead inductances, they may not be as effective at high frequencies.

4.4.4 Fully Differential Circuits

A circuit technique to reduce the effects of parasitic inductances on the power supplies is to use fully differential schemes effective in minimizing the transient current. Figure 4.15 shows an idealized CMOS differential amplifier which is biased by a constant current source and has two inputs and two outputs. As is clear, in steady state conditions, the current flowing from V_{DD} to V_{SS} is constant (i.e., I). If there is a transition, for example $IN+$ rises abruptly from low to high and $IN-$ falls the other way, $OUT+$ is charged toward a high level and $OUT-$ is discharged to a low level. During the transient, current flowing from V_{DD} to V_{SS} is equal to the sum of I and the currents in C1+ and C1-;

$$I_{DD} = I_{SS} = I + C_{1+}\frac{dv_{OUT+}}{dt} + C_{1-}\frac{dv_{OUT-}}{dt}. \tag{4.7}$$

Thus, for a symmetric circuit where $C_+ = C_-$, the supply current remains at I because the last two terms cancel out. When the load capacitors are external to the chip (in case of output drivers), only I_{DD} satisfies Eq. (4.7) and $I_{SS} = I$. Therefore, if a switch is designed fully differentially

throughout the signal paths, the effects of power supply inductances are totally eliminated.

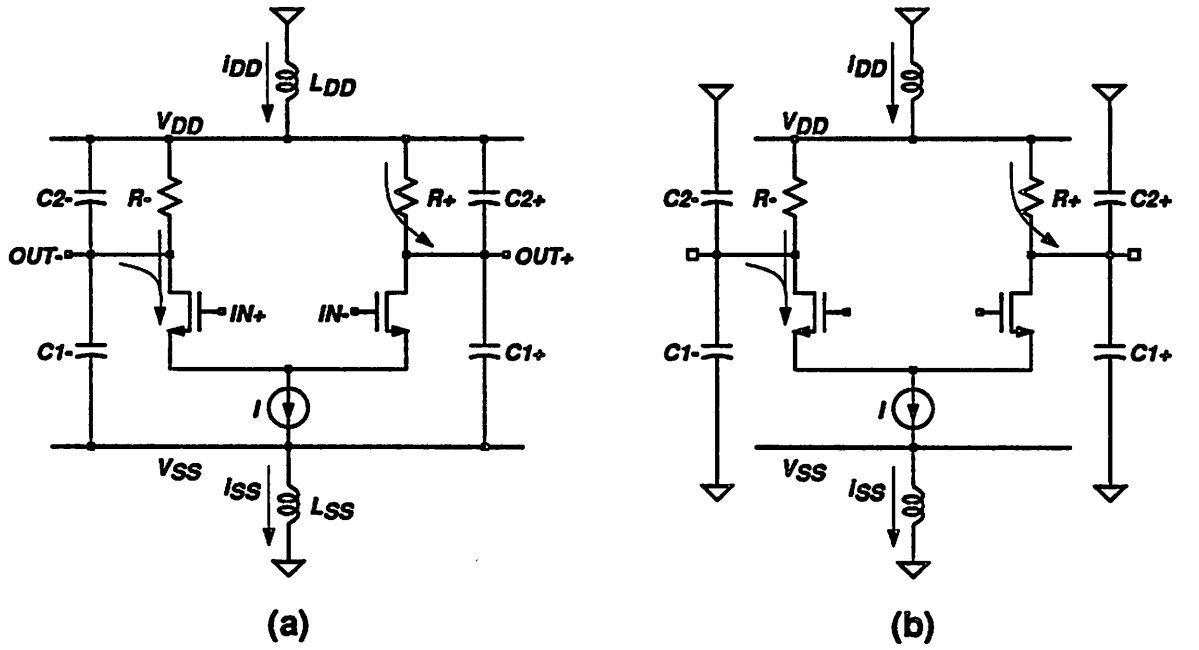


Figure 4.15 Idealized CMOS differential amplifier driving (a) on-chip capacitive loads or (b) external loads.

But the cancellation may not be perfect in actual CMOS circuits because capacitances are voltage dependent. Additionally, when the load resistance is nonlinear like the one of a diode-connected MOSFET load, the cancellation may get worse. However, this technique is still superior to other techniques in minimizing the power supply fluctuation.

One major disadvantage of this technique is that it doubles the number of input/output nodes and needs more area and design effort for crosspoint matrix formation and signal routing. Subsequently, it will increase the on-chip parasitics and degrade the switch bandwidth. Note that, in general, a differential amplifier is already inferior to a simple CMOS inverter in bandwidth.

4.4.5 Constant Current Steering

Another circuit technique that reduces both power supply current transient and signal routing

complexity is single-ended, constant current steering. This technique can eliminate the variation of power supply current, like the fully differential technique.

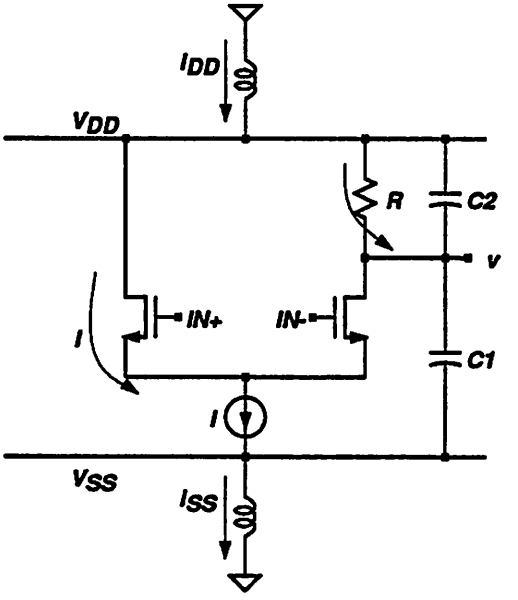


Figure 4.16 Single-ended, differential amplifier.

Let us consider first a single-ended, differential amplifier shown in Figure 4.16. This circuit results from the fully differential amplifier in Figure 4.15 by removing one load. As before, the steady-state supply current is I. But during transients, the supply current is

$$i_{DD} = i_{SS} = I + C_1 \frac{dv}{dt}, \tag{4.8}$$

which is not constant nor equal to I. Hence, this simple, constant current steering circuit makes the supply current to vary, although the variation would be negligible if C_1 is much smaller than C_2.

Next, consider from a simple digital inverter. As indicated in Eqs. (4.5) and (4.6), during a transient, the power supply current for an inverter is not constant but proportional to the transistor current that varies with time. But if the circuit is modified so that constant currents flow through the transistors, the supply current may be kept unchanging as these equations suggest.

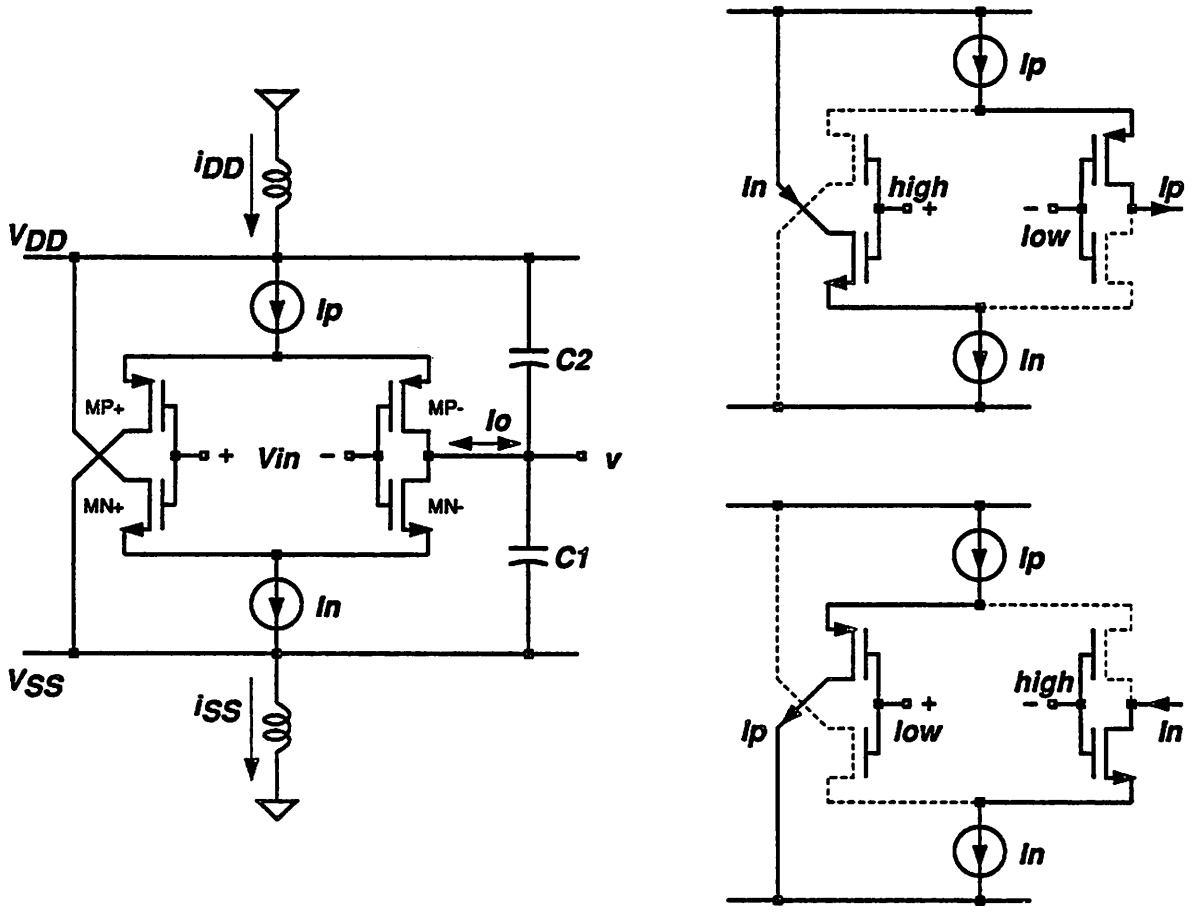


Figure 4.17 Single-ended, complementary, constant current steering circuit.

The constant current steering circuit shown in Figure 4.17 realizes the above concept. It steers two currents, I_n and I_p, to either the output or the power supply, depending on the differential voltage across the two inputs, V_{in}. If V_{in} is positive, MN+ and MP- are ON, I_n flows to V_{DD}, and I_p is steered to the output. And if V_{in} is negative, I_n is switched to the output through MN- and I_p returns to V_{SS} via MP+. Therefore, the output current is constant instead of being dependent on the output voltage like the inverter. Although unrealistic, assuming only capacitive loads at the output to compare with the inverter, the supply currents during transition for a negative V_{in} are

$$i_{DD} = I_p + C_2 \frac{d(V_{DD} - v)}{dt} \tag{4.9}$$

and

$$i_{SS} = I_p + I_n + C_1 \frac{dv}{dt}. \tag{4.10}$$

Since i_{DD} must be equal to i_{SS}, the following relationship holds;

$$(C_1 + C_2) \frac{dv}{dt} = -I_n. \tag{4.11}$$

Hence, the supply current becomes

$$i_{DD} = i_{SS} = I_p + \frac{C_2}{C_1 + C_2} I_n, \tag{4.12}$$

which is constant.

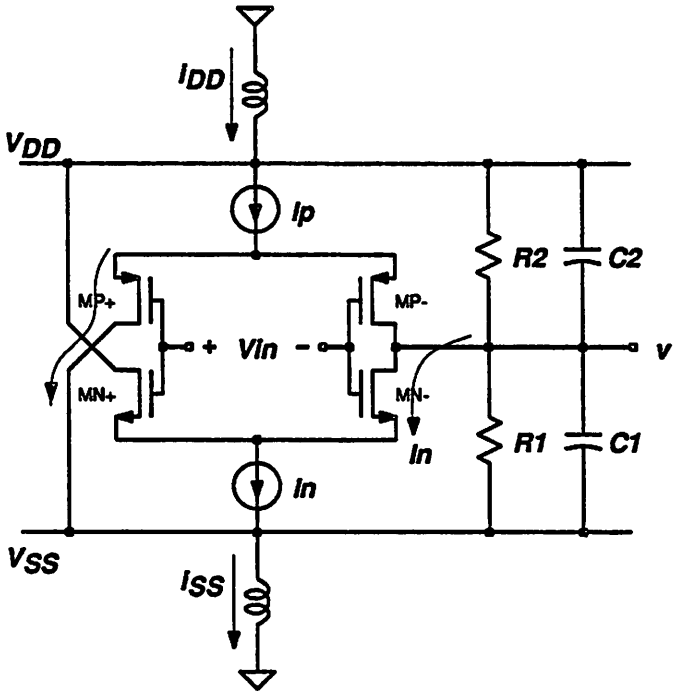


Figure 4.18 Single-ended, complementary, constant current steering circuit with resistive loads.

However in reality, because the current sources are implemented using MOSFETs in the saturation region, to keep the current constant, the output voltages in steady states must be limited via resistive loads. Figure 4.18 shows the current steering circuit with two resistive loads.

The steered current at the output is then converted into a voltage by the two resistors. In essence, this circuit is a complementary version of the circuit in Figure 4.16 and advantageous for symmetric swing applications. Like before, during transition for a negative input, the supply currents are expressed as

$$i_{DD} = I_p + \frac{V_{DD} - v}{R_2} + C_2 \frac{d(V_{DD} - v)}{dt} \tag{4.13}$$

and

$$i_{SS} = I_p + I_n + \frac{v}{R_1} + C_1 \frac{dv}{dt}. \tag{4.14}$$

From the requirement that $i_{DD} = i_{SS}$,

$$\left(\frac{1}{R_1} + \frac{1}{R_2} \right) v + (C_1 + C_2) \frac{dv}{dt} = \frac{V_{DD}}{R_2} - I_n. \tag{4.15}$$

Using this and assuming that $C_1 R_1 = C_2 R_2$, the supply current simplifies and proves to be constant again;

$$i_{DD} = i_{SS} = I_p + \frac{C_2}{C_1 + C_2} I_n + \frac{V_{DD}}{R_1 + R_2}. \tag{4.16}$$

A similar derivation for the corresponding steady state results in the exactly same supply current as this one. Therefore, the supply current for this particular input condition stays constant throughout the transient and steady states.

For a positive V_{in}, the supply current becomes another constant;

$$i_{DD} = i_{SS} = I_n + \frac{C_1}{C_1 + C_2} I_p + \frac{V_{DD}}{R_1 + R_2}. \tag{4.17}$$

This value can be different from the one in Eq. (4.16). But if the two are equal, the supply current of this circuit is invariant regardless of the input polarity. The condition for this is

$$\frac{I_n}{I_p} = \frac{C_2}{C_1}. \tag{4.18}$$

The assumption used in the above derivations is quite reasonable in symmetric designs. In fact, it is natural to design with $I_n = I_p$, $R_1 = R_2$, and $C_1 = C_2$. However in CMOS, the equal complementary currents and equal resistances suggest that p-MOSFETs are 2 to 3 times larger than n-MOSFETs and, thus, C_2 is larger than C_1 by approximately the same factor. In addition, as discussed before, the capacitances are voltage dependent. Due to these capacitance errors, the power supply current may have some transient component and this circuit is not so perfect as the fully differential circuit. But, constant current steering circuits are quite effective in reducing the supply current fluctuation.

4.4.6 Reduction of Voltage Swing

Another simple technique is to limit voltage swing at critical nodes for successive reduction of parasitics and transient currents. This concept is illustrated in Figures 4.19 and 4.20. One of the critical nodes in a switch may be the output node of the crosspoint matrix because it would have a large capacitive load. If the high-speed path is designed with conventional digital CMOS circuits, node voltages swing from V_{SS} to V_{DD} in a time less than the bit period (Figure 4.19). Because average drive current from a crosspoint cell is proportional to the voltage swing for a fixed transition time (see Eq. (4.1)), a large current is required from each cell. To provide this current, devices in the cell need to be large, which impose large parasitic capacitances on both the input and output nodes. Since the input node of the matrix also has a large swing equal to the supply voltage, the input buffer should source a large current and its size must be large. Therefore, the conventional design increases transient current and parasitic capacitances.

Now, let us consider the design in Figure 4.20, in which a voltage swing limiter is simply added at each output of the matrix. Assuming initially that the capacitance is unchanged, because the voltage swing is small compared to the one in Figure 4.19, the drive current from and device sizes in each crosspoint cell can be smaller. The result is a reduction of capacitance contribution to both input and output nodes. This smaller output capacitance further reduces the cell currents

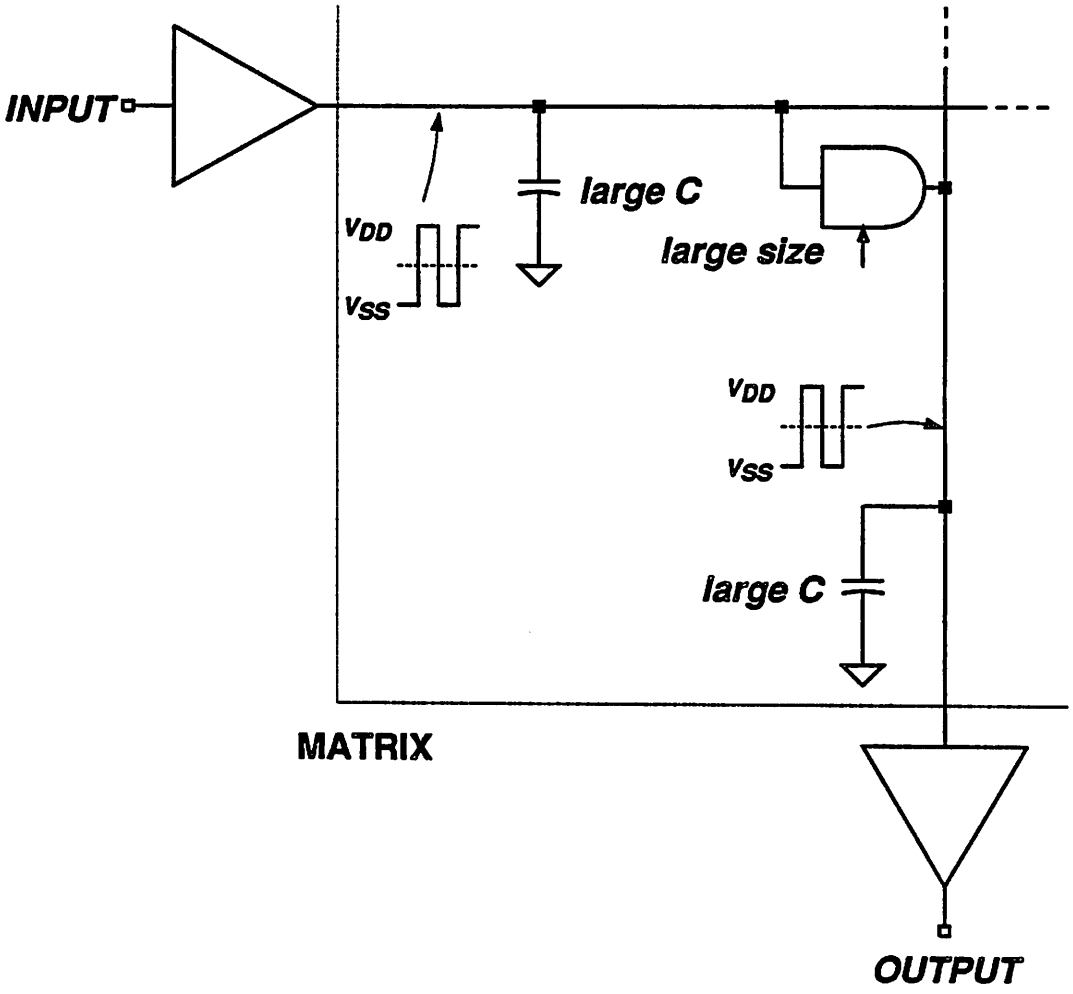


Figure 4.19 Conventional digital CMOS design.

and parasitic capacitances. Hence, a smaller size input buffer is needed to drive a lower capacitive load at the input node. For example, if the voltage swing is reduced by a factor of 10, device sizes and dynamic currents in the crosspoint cell and input buffer as well as capacitances in the matrix are decreased by approximately 10 times. In short, this simple technique simultaneously reduces the supply current transients and the parasitic capacitances.

Voltage swing limiters can be either static or dynamic circuits. Although dynamic circuits using various pulses would allow better control of voltage swing without dissipating static power, they require a much higher technology bandwidth than the switch bandwidth because the pulse

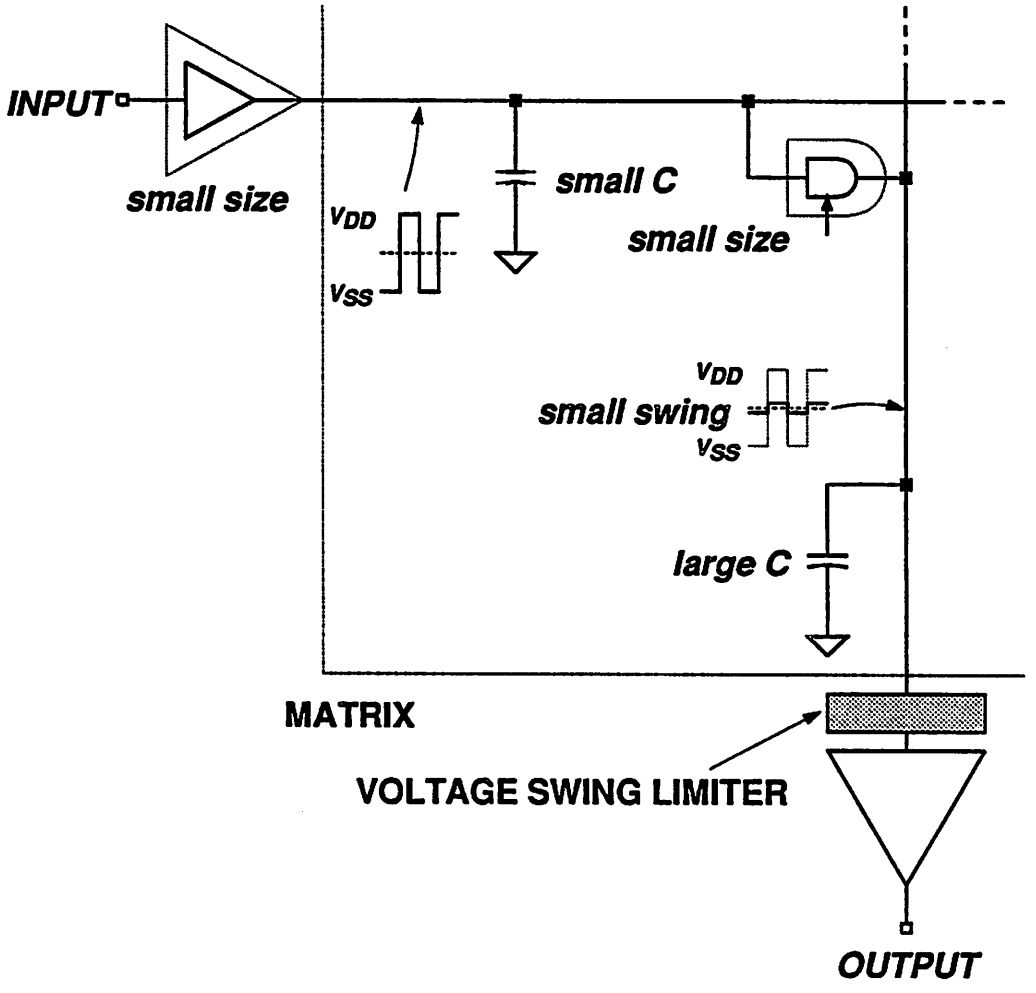


Figure 4.20 Reduced voltage swing design.

periods must be smaller than the bit period. Static circuits are advantageous in this sense, but may consume more static power.

Reduction of voltage swing is primarily based on lowering output load resistance of a gain stage such as an inverter. A simple technique is then to add resistive loads at the output as shown in Figure 4.21. The resistance can be linear or nonlinear (e.g., diode). In this circuit, the voltage swing is determined as

$$V_S = (I_n + I_p) R_1 \| R_2, \tag{4.19}$$

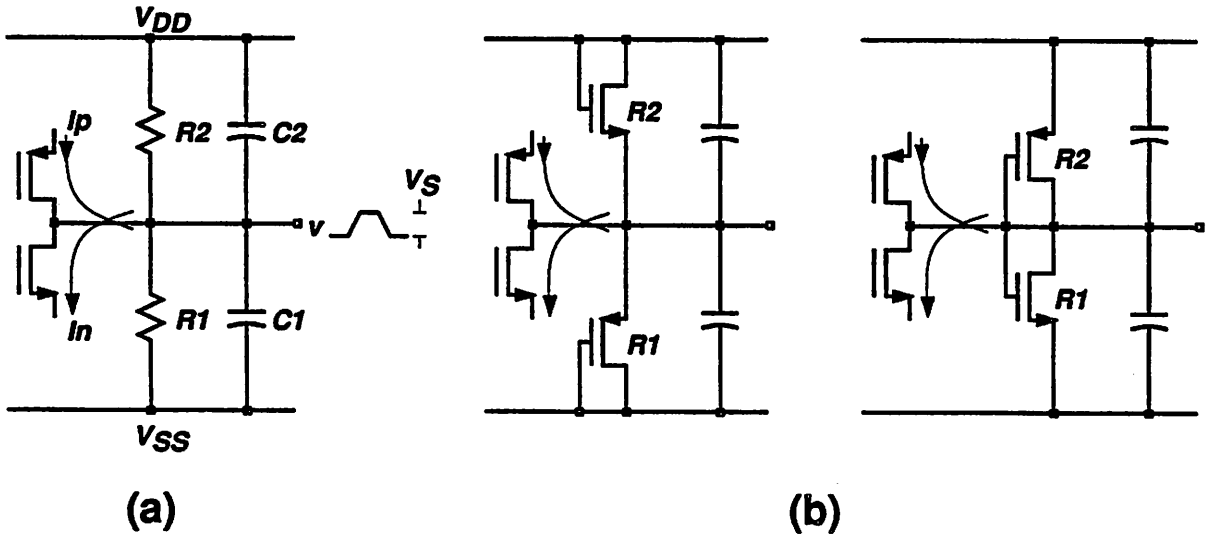


Figure 4.21 Voltage swing is reduced via (a) linear or (b) nonlinear resistors.

where I_n and I_p represent currents flowing in and out of a logic gate, respectively, depending on the logic status.

The currents may be constant during transients like the current steering case or varying like the simple inverter case. But note that the voltage swing reduction helps to minimize the current variation of the simple inverter. This is because the drain voltage varies only V_S for entire transitions. Since the MOSFETs very likely will be in the linear region, the current change will be approximately proportional to $(V_{GS} - V_T - V_{DS})$, where V_{DS} is the average value of drain-to-source voltage. Therefore, from Eqs. (4.5) and (4.6), the power supply current, even for the simple inverter, becomes roughly constant during transitions. This is an additional advantage of the voltage swing reduction technique.

One drawback of using real resistors is the difficulty in implementing them in CMOS and matching the average output level and the logic threshold of the next stage. The matching is helpful for symmetric signal propagation. Resistors realized from an inverter that has an identical logic threshold with the next stage and whose input and output are tied together are better,

although the matching may not be perfect. In addition, because these resistors are diode-connected MOSFETs, more current is available for discharging capacitances at the initial phase of a transition. Noting that the resistance of a MOSFET diode is inversely proportional to $V_{GS} - V_T$, if the output falls from high to low, initially R_2 is larger and R_1 is smaller than the nominal values.

A major disadvantage of this technique is a large static power consumed in steady states. For the above circuit when the output is low, the static power consumption is as large as

$$P_S = \frac{V_{DD}^2 + V_{DD} R_1 I_n}{R_1 + R_2}. \tag{4.20}$$

However, this power increase is offset by the decrease in dynamic power dissipation given by

$$P_D = (C_1 + C_2) V_S^2 f. \tag{4.21}$$

Now, let us compare the total power consumption with the conventional design, under the assumption that $R_1 = R_2$. The total power in this circuit is the sum of average static power and dynamic power, and that in the ordinary CMOS design is dynamic only. Using Eq. (4.19), the ratio becomes

$$\frac{1}{2R(C_1 + C_2)f}\left(1 + \frac{V_S}{V_{DD}}\right) + \frac{V_S^2}{V_{DD}^2} \tag{4.22}$$

that decreases as either f gets higher or V_S is reduced. Note that the ratio is about $\frac{\pi}{2}$ at the corner frequency of this circuit, i.e., $\frac{1}{\pi R(C_1 + C_2)}$.

Another issue of this reduced swing circuit is noise immunity. In crosspoint switches, primary source of noise is capacitive coupling from other nodes. If resistors are added at the output of the ordinary gate without reducing the device sizes, the corner frequency would increase. But remembering that main purpose of the swing reduction at highly capacitive nodes is to reduce the device sizes while maintaining the corner frequency, equivalent output resistance of the low-

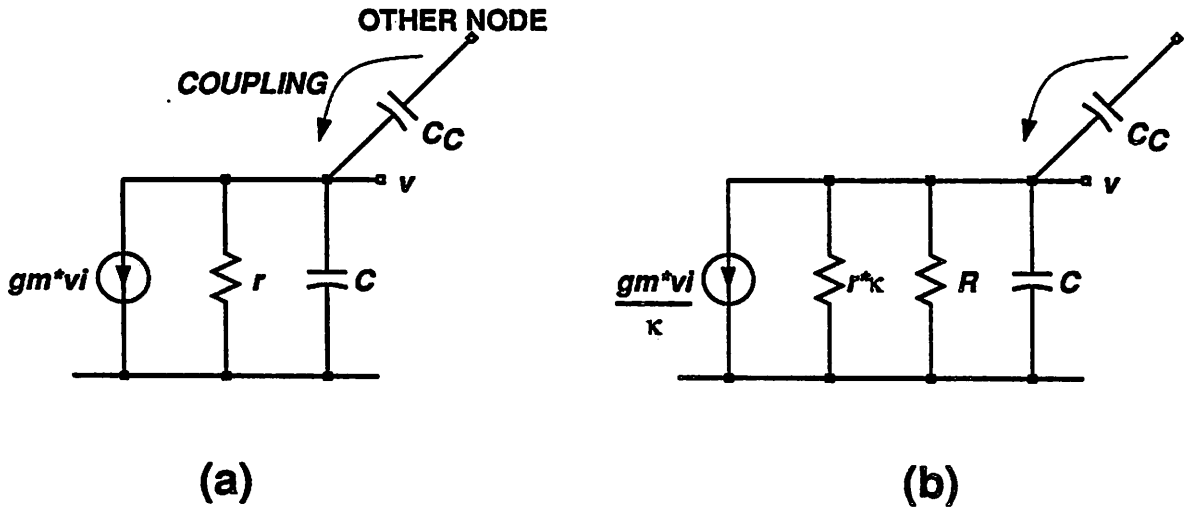


Figure 4.22 Simplified models for coupling in (a) high-swing design and (b) low-swing design.

swing circuit must be very close to the one in high-swing circuit (Figure 4.22). That is, for a size reduction by κ, $r = (\kappa r) \| R$. Therefore, the degree of coupling from other nodes remains unchanged if C_C is constant. Of course, when C_C comes mostly from the gate-drain overlap, the degree of coupling will decrease with the size reduction. Also, since the output signal levels in the matrix and power supply noise level are similarly reduced in this design, the *signal-to-noise ratio* (SNR) may be maintained. However, there are signals still having a large voltage swing such as inputs to the matrix and controls. If the coupling is from these nodes, the ratio will be lowered. In addition, if the swing is too small, the SNR will be degraded due to increasing portion of device related noises.

Voltage swing can also be reduced by using extra bias sources and hard-limiting diodes as shown in Figure 4.23. Although the swing is exactly controllable, these methods may suffer from the level matching problem mentioned above and layout complexity. Especially in (a), the extra bias sources must be routed to every crosspoint cell. The method in (b) may not be practical in CMOS, due to the lack of good diodes. Diode-connected MOSFETs can be used, but they need to be large and may have severe body effect. Instead of being used as diodes, the MOSFETs can

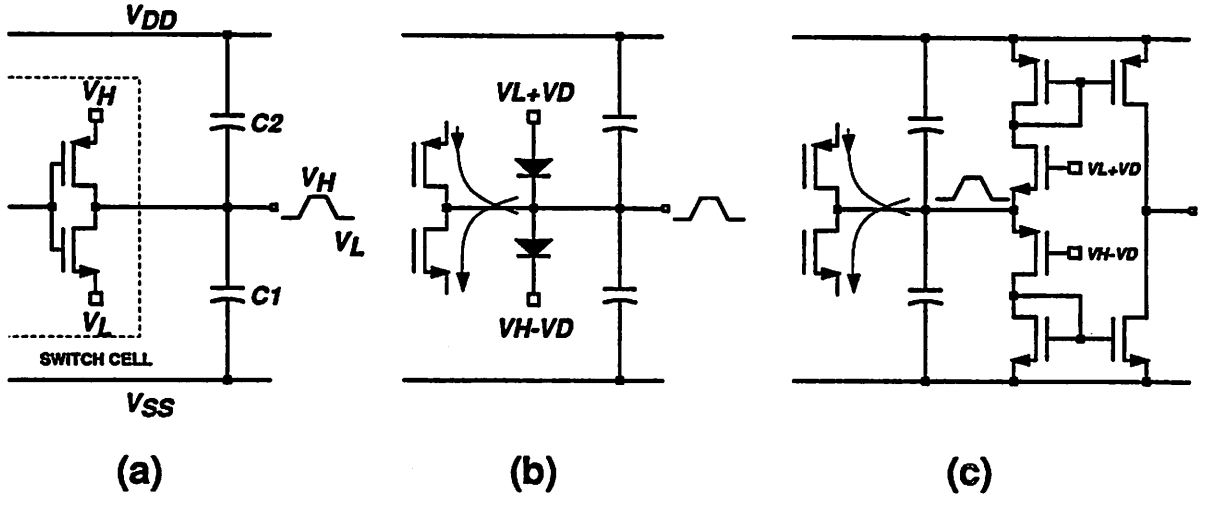


Figure 4.23 Swing reduction using extra bias sources and diodes.

be utilized for a cascode amplifier like the one in (c). This circuit has dual function: the swing limitation and amplification. However, like (b), this circuit is more sensitive to couplings due to higher impedance at the output node during transitions.

4.4.7 Transresistance Amplifier

A suitable technique to limit voltage swing and match the center level with the logic threshold of the next stage is to convert the next stage into a transresistance amplifier that is composed of a wideband amplifier and a feedback resistor as illustrated in Figure 4.24. Basically, this circuit senses an input current and converts it into a voltage through the feedback resistor. Then, voltage swing at the input node is very small because the input is at the virtual ground when the amplifier gain is large. The swing is represented by Eq. (4.19) with $R_1 \| R_2$ replaced by the effective input resistance of this circuit.

Most common CMOS amplifiers are differential amplifiers like the one in Figure 4.15. Their gains are made high by using low drain-saturation voltages [46]. But, despite high gains, they are generally not suitable for this application due to insufficient bandwidths.

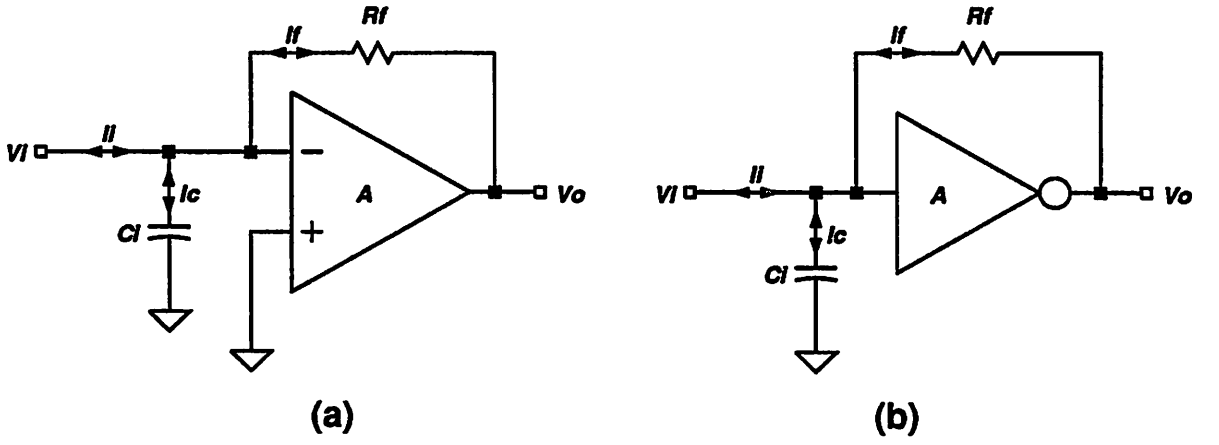


Figure 4.24 Swing reduction using (a) transresistance amplifier and (b) its inverter equivalent.

In digital CMOS technologies, an inverter biased at the logic threshold realizes a good wideband amplifier, even though the gain is low. Figure 4.25 shows the DC input-output transfer curve of a 1.4 μm channel length inverter. In the steep transition region, the two MOSFETs are in the saturation region and the DC gain is about 12. The corner frequency of this inverter is as high as 250 MHz. A higher gain can be achieved by using longer channel devices, but the resulting bandwidth gets narrower rapidly. Since an inverter is biased automatically at the logic threshold by connecting the input and output using a resistor, a wideband transresistance amplifier can be obtained simply by adding a feedback resistor around an inverter (Figure 4.24 (b)). Note also that, because the matrix output is usually connected to an inverter for buffering and this inverter may be converted into a transresistance amplifier, this technique allows combination of voltage swing reduction and amplification in one stage.

Feedback resistors can be made using resistive layers in CMOS technologies, such as polysilicon and diffusion layers. Although this method looks straightforward, it takes large area to get reasonable resistances, and adds substantial parasitic capacitances. In addition, the absolute values vary a lot following the processes and this variation is independent of the variation of currents that flow through the resistors.

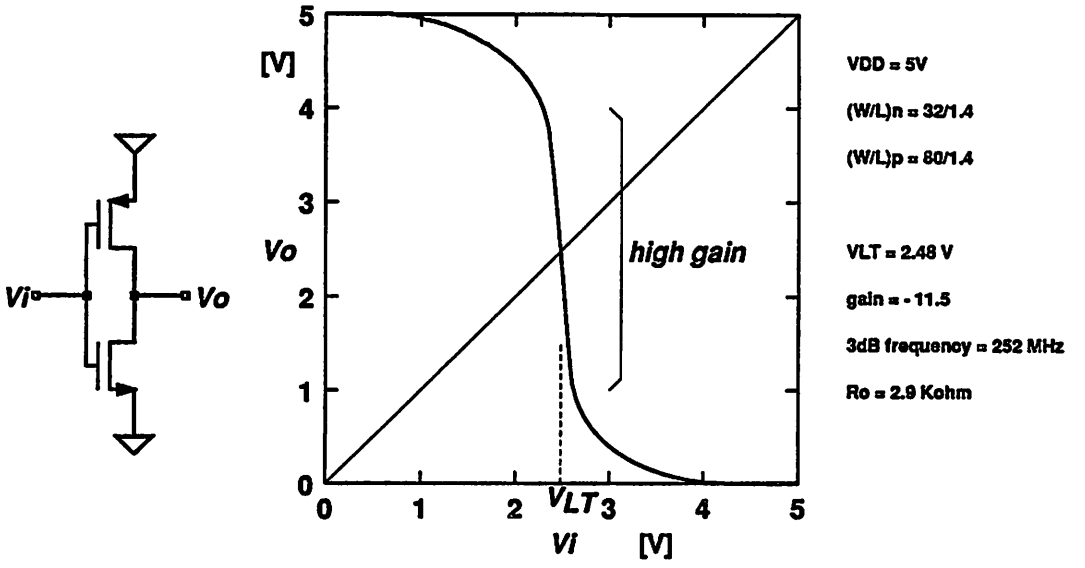


Figure 4.25 DC transfer curve of an inverter.

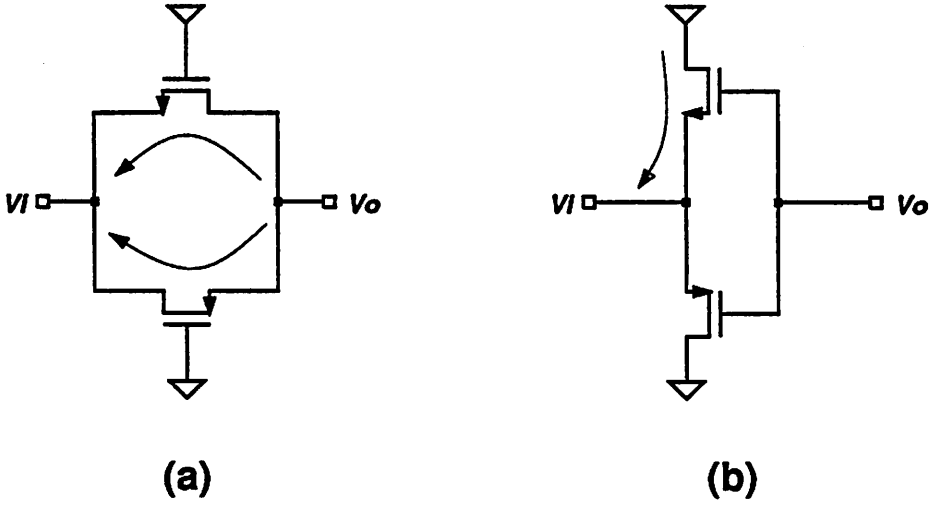


Figure 4.26 CMOS resistors: (a) linear and (b) nonlinear.

The resistors can also be implemented with MOSFETs (Figure 4.26). This design is advantageous over others made of polysilicon or diffusion, in that the current variation may offset the resistance variation, producing a relatively constant output swing. Furthermore, the gain and frequency response can be optimized by controlling aspect ratios of the transistors.

One simple method is to utilize CMOS transmission gates. Because the gate of n-MOSFET is tied to V_{DD} and that of p-MOSFET is connected to V_{SS}, n-MOSFET will be ON as long as either input or output potential is below $V_{DD} - V_{Tn}$, and p-MOSFET is ON if either one is above $V_{SS} + |V_{Tp}|$. The resistance of this transmission gate does not vary much over a large range of signal swing (normally, equal to the power supply voltage). And the linearity would improve if the body effects on both transistors are negligible. Unfortunately, there is no differentiation between source and drain for these MOSFETs. So, connecting the bodies with sources to eliminate the body biases will not work.

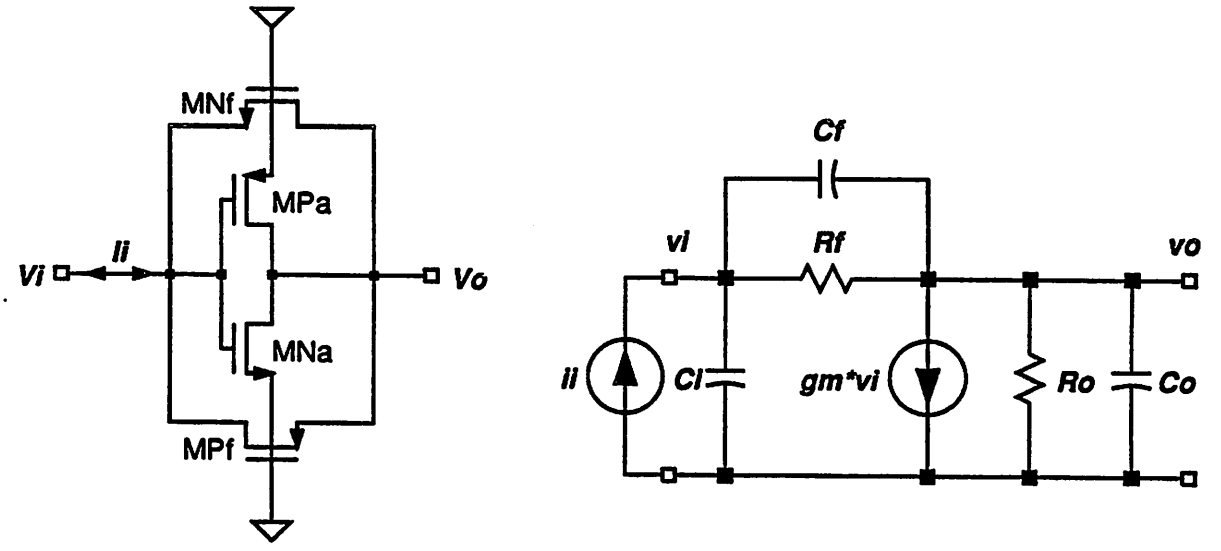


Figure 4.27 Transresistance amplifier using inverter and transmission gate, and its small-signal model.

Another way is to make a nonlinear resistor like the one in Figure 4.26 (b). This resistor acts as hard-limiting diodes as before (see Figure 4.23 (b) and (c)) when there is a current. Note that the current does not flow to the output node but to the power supply. However, a voltage still develops between the input and output, i.e., the gate and source of one transistor, according to the current flowing through. The voltage is approximately proportional to the square-root of the current. Hence, when this resistor is used in a feedback loop, the forward and feedback paths can

be controlled separately for optimum performance. A major disadvantage is that, for zero current, the output potential is not determined and may oscillate.

A CMOS transresistance amplifier or current-to-voltage converter composed of an inverter and a transmission gate is shown in Figure 4.27. In steady states, the input current I_i (I_n or I_p) must flow through the two feedback transistors (MNf and MPf) to the output node, because the gate of a MOSFET cannot conduct direct current. It, then, returns to the supply through either of the inverter transistors (MNa and MPa). Since the voltage drop across the feedback resistor, I_iR_f, should also drop between the input and output, MNa and MPa will have different bias voltages. These bias voltages are set at a point where the resulting unbalance of inverter current matches the input current. Note that, at the logic threshold of the inverter, MNa and MPa are both in the saturation region, and their drain currents are primarily proportional to $(V_{GS} - V_T)^2$. Therefore, if the input current is small, the input voltage deviates little from the logic threshold.

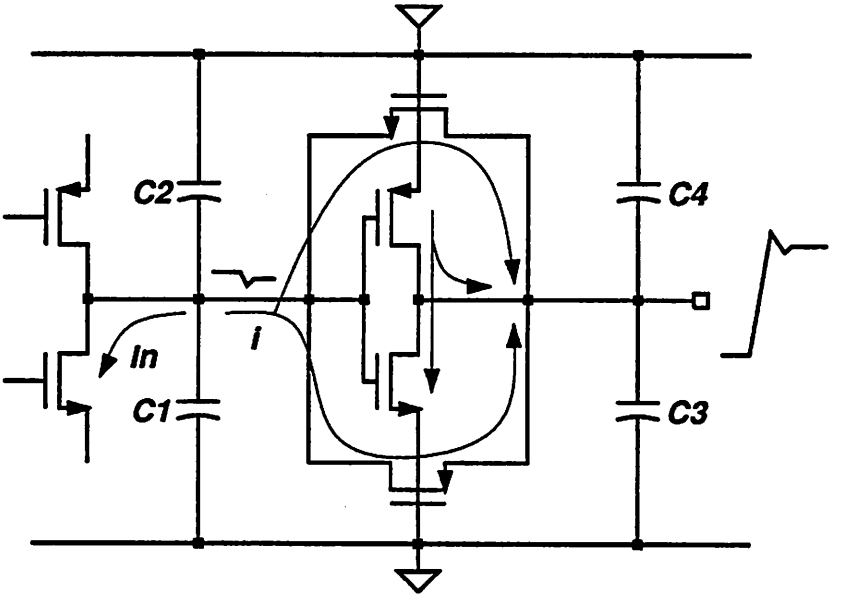


Figure 4.28 Currents during a transient in the transresistance amplifier.

When the input current changes status, say, from I_p to I_n, the input node is discharged ini-

tially by a current $I_p + I_n$ because the output voltage does not change quickly and I_p still flows through R_f (Figure 4.28). Then the input voltage falls and the output voltage starts to rise, which reduces the voltage drop across R_f and current through it. The discharge continues as long as the feedback current into the input node is smaller than I_n. Once the two currents match, discharge stops. But due to the delay, the input node may be charged again by the excessive feedback current. This causes decaying oscillation and the circuit finally settles in a steady state. As is evident, if the feedback action is fast, overshoot of the input and output decreases. Roughly, larger R_f and output capacitance make the action slower. A larger R_f also increases the output swing. But if increased too much, the output is cramped and the input swing increases. In this case, one of the MOSFETs in the inverter gets in the linear region and the gain drops rapidly.

The resistance of the transmission gate when the two transistors are in the linear region, i.e., both V_i and V_o lie in between $V_{DD} - V_{Tnf}$ and $V_{SS} + |V_{Tpf}|$, is obtained by summing the two current components and differentiating it with respect to $V_i - V_o$;

$$G_f \equiv \frac{1}{R_f} = k'(\frac{W}{L})(V_{DD} - V_{Tnf} - |V_{Tpf}|). \tag{4.23}$$

Here, the gain factors of both transistors are assumed to be identical to $k'(\frac{W}{L})$, and V_{Tnf} and V_{Tpf} are threshold voltages of MNf and MPf, respectively. This equation suggests that the resistance is constant for a large range of input/output voltages. The resistance becomes dependent on the input/output levels, if one of the transistors is in saturation.

Actual resistance has more complicated behavior, due to the inaccuracy of MOSFET models and the body effects. The simple long-channel model used in this derivation fails as shorter channel devices are used for higher speeds. If resistances required are high (> 5 kohm), they may have to be made with long-channel devices. The input/output voltages impose body biases that affect the device threshold voltages in the increasing direction. So, R_f would be a function of V_i and V_o. However, the functional dependency is weak, because the increase in, say,

V_{Tnf} due to higher input/output voltages is offset by the decrease in $|V_{Tpf}|$. This resistor is also sensitive to the power supply and process variations; k' and V_T's are quite process dependent. Also these parameters are weak functions of the supply voltage and geometry through the body effects and the short-channel effects: the mobility degradation and velocity saturation [47].

But if the sensitivity matches that of the input current, the circuit may be designed to have a relatively constant output swing. As an example, let us consider a simple inverter driving this transresistance amplifier. And assume that k' is constant for different devices and biases, and variation of threshold voltages with respect to the supply voltage is negligible. Since the input node of this amplifier has a very small swing biased at the logic threshold that is usually $\frac{1}{2}V_{DD}$, the driving MOSFETs will be in the linear region. Therefore, if the supply voltage increases, the current rises rapidly through V_{GS} and V_{DS} of the inverter. On the other hand, the resistance R_f decreases according to Eq. (4.23). Then, the voltage drop across the resistor, I_iR_f, becomes

$$I_iR_f = \frac{1}{2}(\frac{W}{L})_r \frac{\frac{3}{4}V_{DD}^2 - V_TV_{DD}}{(V_{DD} - V_{Tnf} - |V_{Tpf}|)} \tag{4.24}$$

where $(\frac{W}{L})_r$ defines the ratio of the driving MOSFET's aspect ratio to the resistor's aspect ratio, and V_T is the threshold voltage of the driver. Note that the drop is independent of k'. Now, the sensitivity of this drop to V_{DD} is

$$S_{V_{DD}}^{I_iR_f} = \frac{\frac{3}{4}V_{DD}^2 - \frac{3}{2}(V_{Tnf} + |V_{Tpf}|)V_{DD} + V_T(V_{Tnf} + |V_{Tpf}|)}{\frac{3}{4}V_{DD}^2 - [V_T + \frac{3}{4}(V_{Tnf} + |V_{Tpf}|)]V_{DD} + V_T(V_{Tnf} + |V_{Tpf}|)}. \tag{4.25}$$

Assuming that all V_T's are very close to $\frac{1}{5}V_{DD}$, the sensitivity becomes about 0.7. Whereas, that of a constant resistor is as large as 2.4. If the current is generated by dividing a reference voltage that is insensitive to the supply variation, by a similar MOSFET resistor, the sensitivity of I_iR_f becomes approximately equal to that of the reference voltage.

The equivalent input resistance of this transresistance amplifier is obtained from the small-signal model as

$$R_{ieq} = \frac{G_o + G_f}{(g_m + G_o)G_f} \simeq \frac{R_f}{A_{vo}}, \tag{4.26}$$

where the inverter gain $\frac{g_m}{G_o}$ has been assumed to be much larger than 1, and $A_{vo} \equiv \frac{g_m}{G_o + G_f}$ is the amplifier gain with the feedback element loading effect. Thus, the voltage swing at the input node represented by Eqs. (4.19) and (4.26) can be made small by increasing the amplifier gain. But the gain cannot be greater than the inverter gain that is already limited as discussed before.

The transfer function derived from the small-signal model suggests that the transimpedance has a low-pass characteristics governed by two poles and one zero. At low frequencies, the transresistance is expressed as

$$R_t \equiv \frac{V_o}{I_i} = \frac{G_f - g_m}{(g_m + G_o)G_f}. \tag{4.27}$$

Because this value must be negative for proper operation, g_m has to be greater than G_f. If $g_m \gg G_f$, $R_t \simeq -R_f$.

The bandwidth of this circuit is practically maximized when the two poles become identical. In this condition, the output transient response is critically damped and the output settles fast without ringing. However, design equations for the critical damping involve a large number of design parameters. As a result, without simplifying assumptions, effects of these parameters may not appear clearly. Since this circuit is used to reduce the voltage swing at a highly capacitive node, C_i that is sum of this capacitance and gate capacitance of the inverter would be very large. On the other hand, C_f that accounts for gate-drain overlap is negligible. C_o that represents output capacitance of the amplifier and input capacitance of the next stage is also large but usually smaller than C_i. Otherwise, the main point of this technique would be lost. The transfer function based on the above assumptions predicts the zero placed at a much higher frequency than the

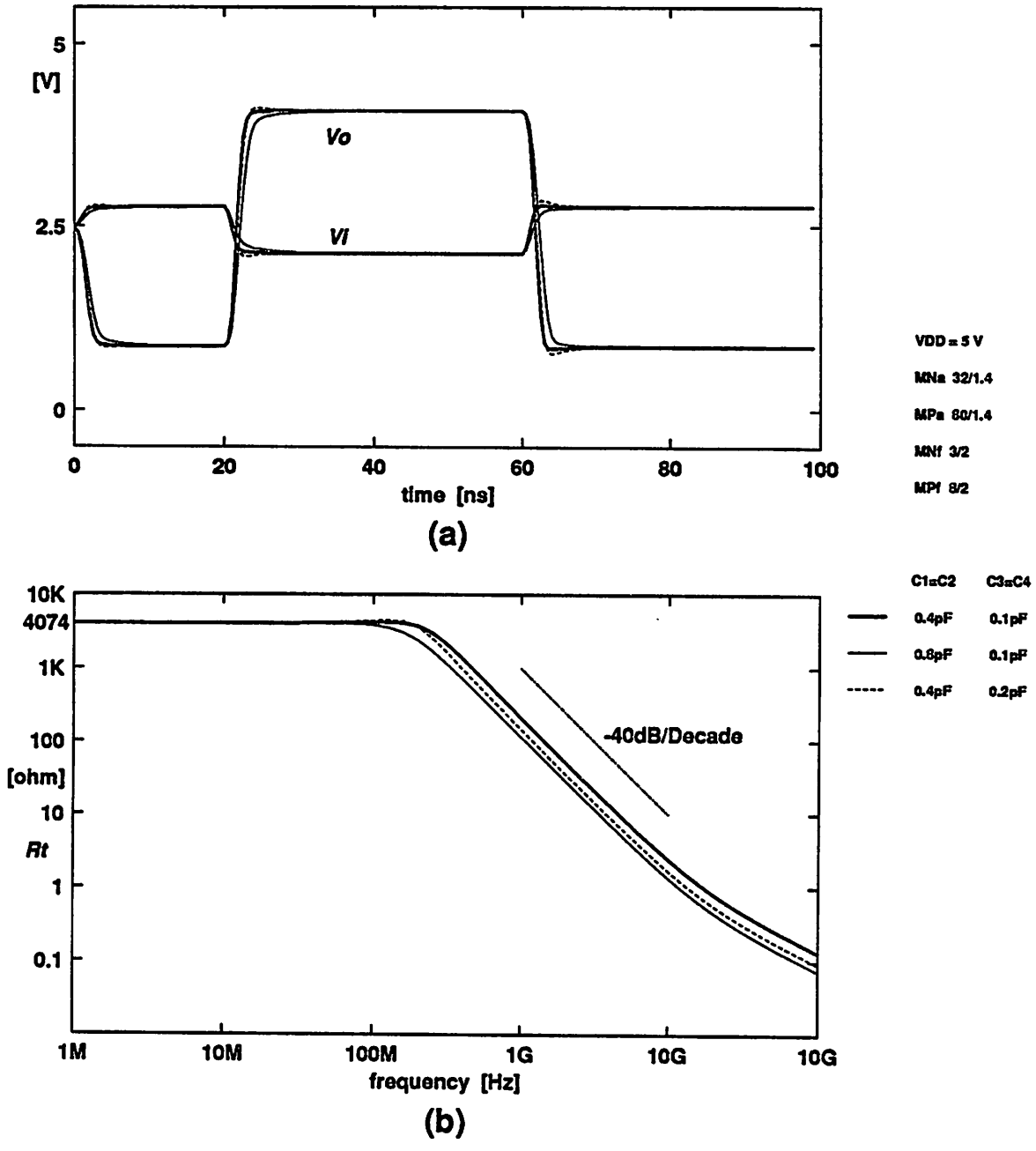


Figure 4.29 (a) Simulated transient response and (b) transfer function of a transresistance amplifier.

poles. Further assuming that $\frac{C_i}{C_f} \gg \frac{g_m}{G_o}$, the condition for critical damping becomes

$$C_i(G_f + G_o)^2 \approx 4C_o g_m G_f. \tag{4.28}$$

Therefore, if $G_f \approx G_o$, the capacitance ratio $\frac{C_i}{C_o}$ should be close to the inverter gain for best performance. The design equation for G_f is then

$$G_f \approx g_m(2\frac{C_o}{C_i} - \frac{G_o}{g_m}) \pm 2g_m[\frac{C_o}{C_i}(\frac{C_o}{C_i} - \frac{G_o}{g_m})]^{1/2}. \quad (4.29)$$

Because G_f must be real, the following condition has to be satisfied:

$$\frac{C_i}{C_o} \leq \frac{g_m}{G_o}. \quad (4.30)$$

Any change in the parameters after being set for critical damping causes the output over or under damping. For example, for a fixed G_f, increase in C_i and G_o makes over damping, while increase in C_o and g_m adds ringing. Also, if G_f is changed to a value between the two in Eq. (4.29), the output will be under damped. The above behavior is also applicable for the input swing. Simulated transient response and transfer function of an example transresistance amplifier are shown in Figure 4.29. This example has a maximum bandwidth of about 250 MHz and verifies the above predictions.

4.4.8 Small-Swing Input and Output

Reduction of the input/output voltage swing is beneficial for the multiple-channel, high-speed switch, as it reduces the supply noise and power consumption. Small voltage swings also lower the probable overshoot of the input/output beyond the on-chip power supply voltage and, in turn, help reduce the possibility of latch-up. Although designing the input/output levels compatible with ECL would be desirable from a system point of view, good output stages are hard to implement in CMOS. A sub-optimum solution may be to design for a small-swing not necessarily matching the ECL levels. As long as the input/output levels have a proper interface, the switches designed this way can be cascaded directly to form a large switching network. Also, if optical transceivers are designed for this interface, they can be simply connected to the switch.

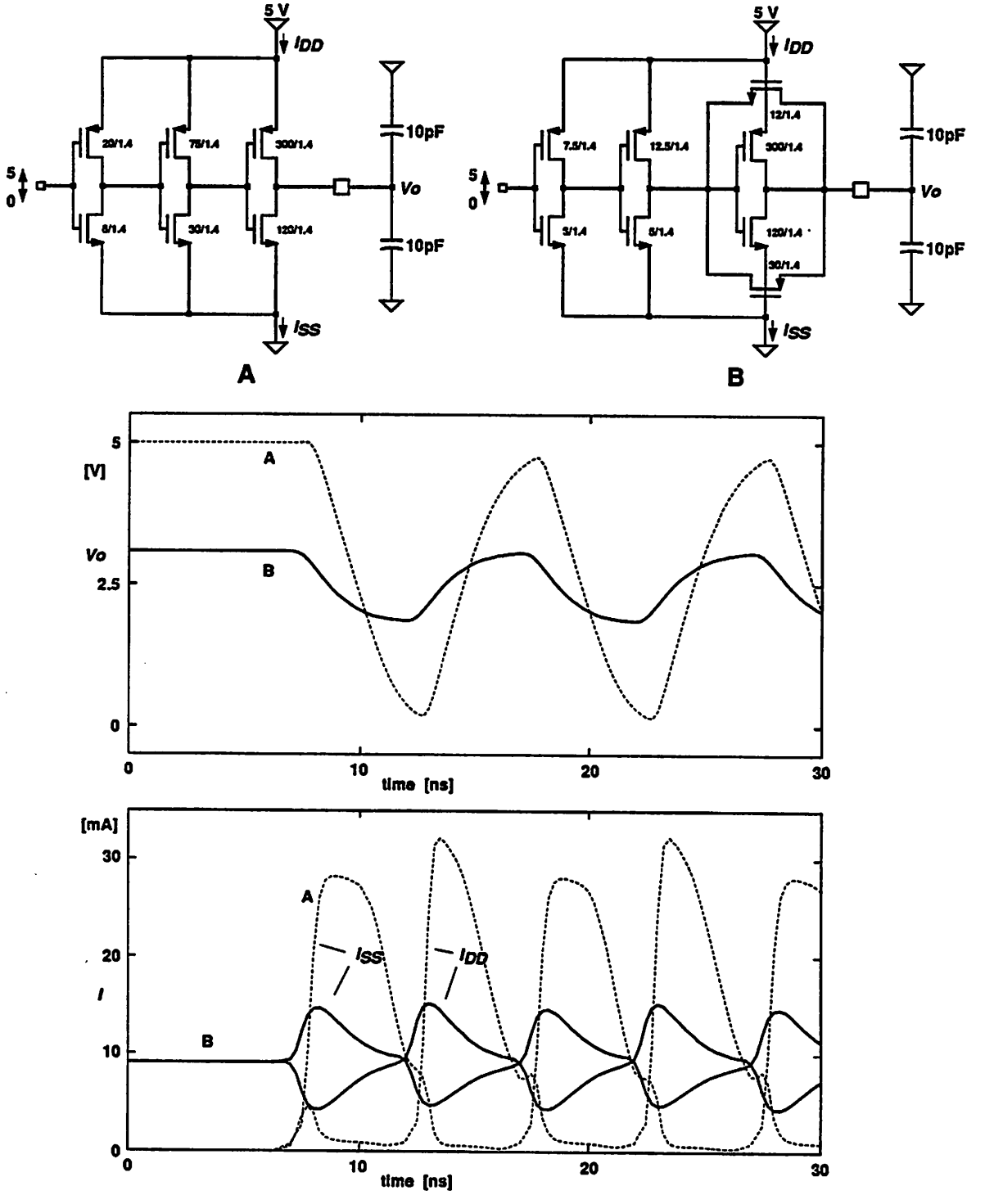


Figure 4.30 Comparison of output stages showing drastic reduction of current transient by reducing output voltage swing.

The input swing may be reduced by using differential stages. In CMOS, the swing can be as low as $\sqrt{2}(V_{GS} - V_T)$ where V_{GS} is the gate bias of the input MOSFET [46]. One way to reduce the output swing is to utilize the transresistance amplifier shown above, differently. Remembering that the amplifier is basically a current-to-voltage converter, output voltage of this circuit can be controlled by the current and the feedback resistance. Note, here, that this circuit is used for output swing reduction, not input swing reduction. As shown in Figure 4.30, this technique reduces the transient current drastically compared to an ordinary digital CMOS circuit for the same output load. In this example, a voltage swing reduction by about a factor of 3.6 decreases the current swing by 2.8 and maximum di/dt by 6.5 times. But this output stage may suffer from slow settling, since the off-chip capacitance load at the output is usually larger than the capacitance at the input.

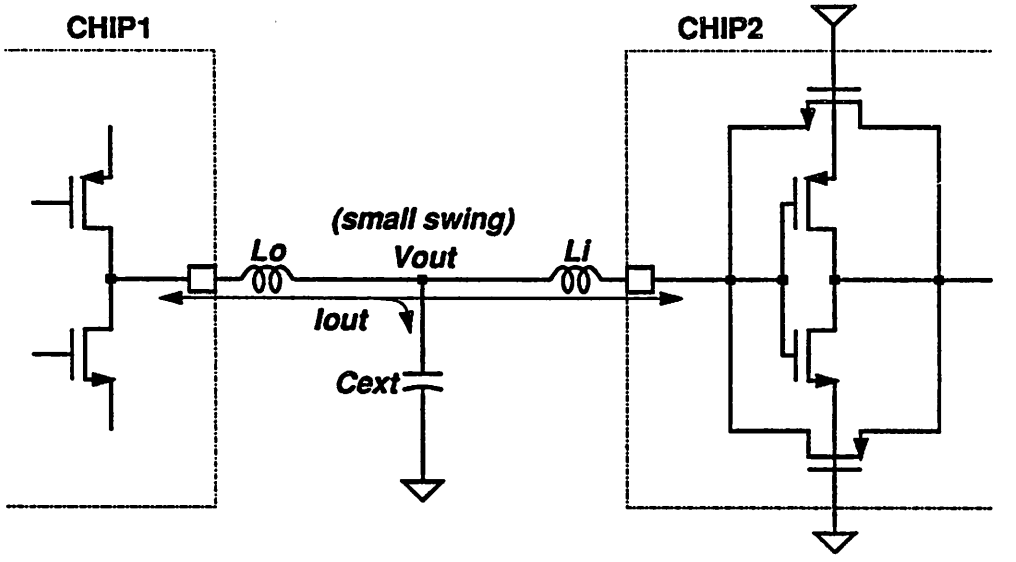


Figure 4.31 Current driving output and current sensing input for output voltage swing reduction.

Another technique to limit the output voltage swing is to design the output circuit to source or sink current and the input circuit to sense this current. An example is illustrated in Figure 4.31, where input stage of the switch is a transresistance amplifier that converts current from the

output driver while cramping the output swing. Because the external capacitance is large, this is a natural application area of the transresistance amplifier. But a large current swing that is needed to charge the off-chip capacitance would be harmful, due to packaging inductances between the output and input.

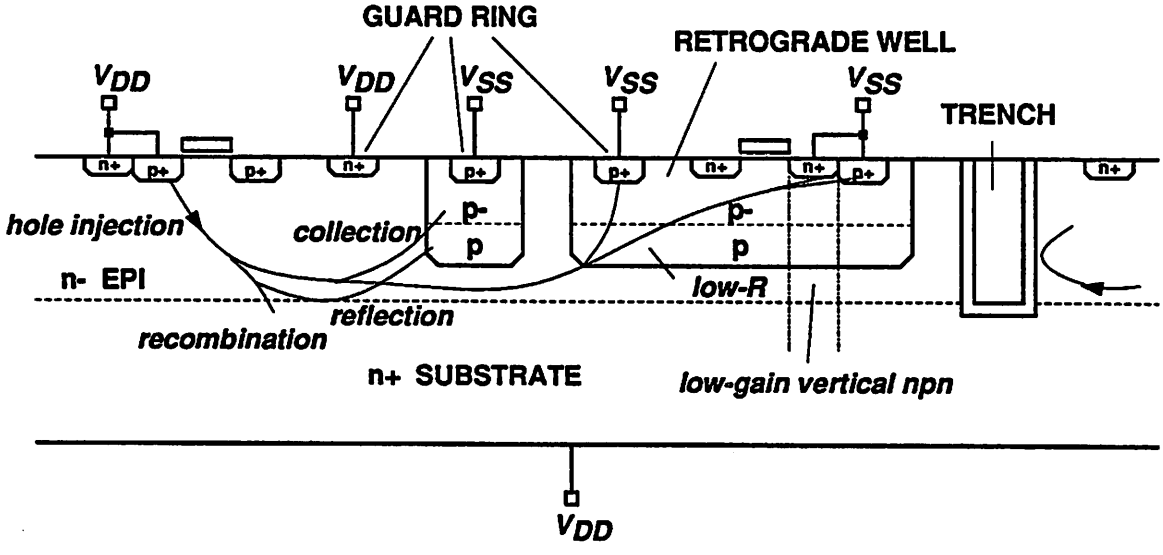


Figure 4.32 Latch-up minimization techniques.

4.4.9 Latch-up Protection

Increasing possibility of the CMOS latch-up due to the power supply and internal voltage fluctuation can only be reduced by providing extensive guard-ring diffusions around the devices, more substrate and well contacts, heavier input/output protection, and substrate contact ring at the chip periphery, for a given technology [48]. If there is a freedom to choose technologies, the epitaxial substrate [49], retrograde well [50], or trench isolation [51] technologies are advantageous. As depicted in Figure 4.32, these techniques either lower body resistances, reduce parasitic bipolar gain, or collect or block carriers before they trigger latch-up.

CHAPTER 5

PROTOTYPE CMOS SWITCH

This chapter describes the design and implementation of a prototype, 16 × 16 switch chip with a capacity of 100 to 250 Mb/s and an operating voltage of 5 V, in CMOS VLSI technologies.

As has been discussed in the previous two chapters, CMOS technologies with minimum channel lengths below 2 μm are attractive for 200 Mb/s electronic switching or higher. But practical implementation of switching components in these technologies would be complicated by the high-speed requirement, their digital nature, and the increasingly apparent parasitics. To satisfy the speed requirement while minimizing waveform degradation, techniques that fully utilize the marginal technology bandwidth and reduce the effects of parasitics are in demand. The circuit techniques suggested in Chapter 4 such as voltage swing limitation at highly capacitive nodes and constant current steering are examples of such techniques. The effectiveness of the circuit solutions will be proven by applying them to this prototype switch. Discussed in this chapter are the architecture, circuit details, and measurement results of the experimental chip as well as the future improvements.

5.1 ARCHITECTURE

The prototype architecture of the switch is shown in Figure 5.1. It is based on the matrix type switch architecture discussed in Section 3.4.2 and shown in Figure 3.7. This architecture is chosen primarily because it is good for LAN applications and enables an efficient layout. The chip consists of high-speed paths that include 16 × 16 crosspoints and input/output buffers (shaded in the figure) and of memory and control blocks. Each crosspoint cell in the switch plane

is bit-mapped to the corresponding memory cell in the memory plane.

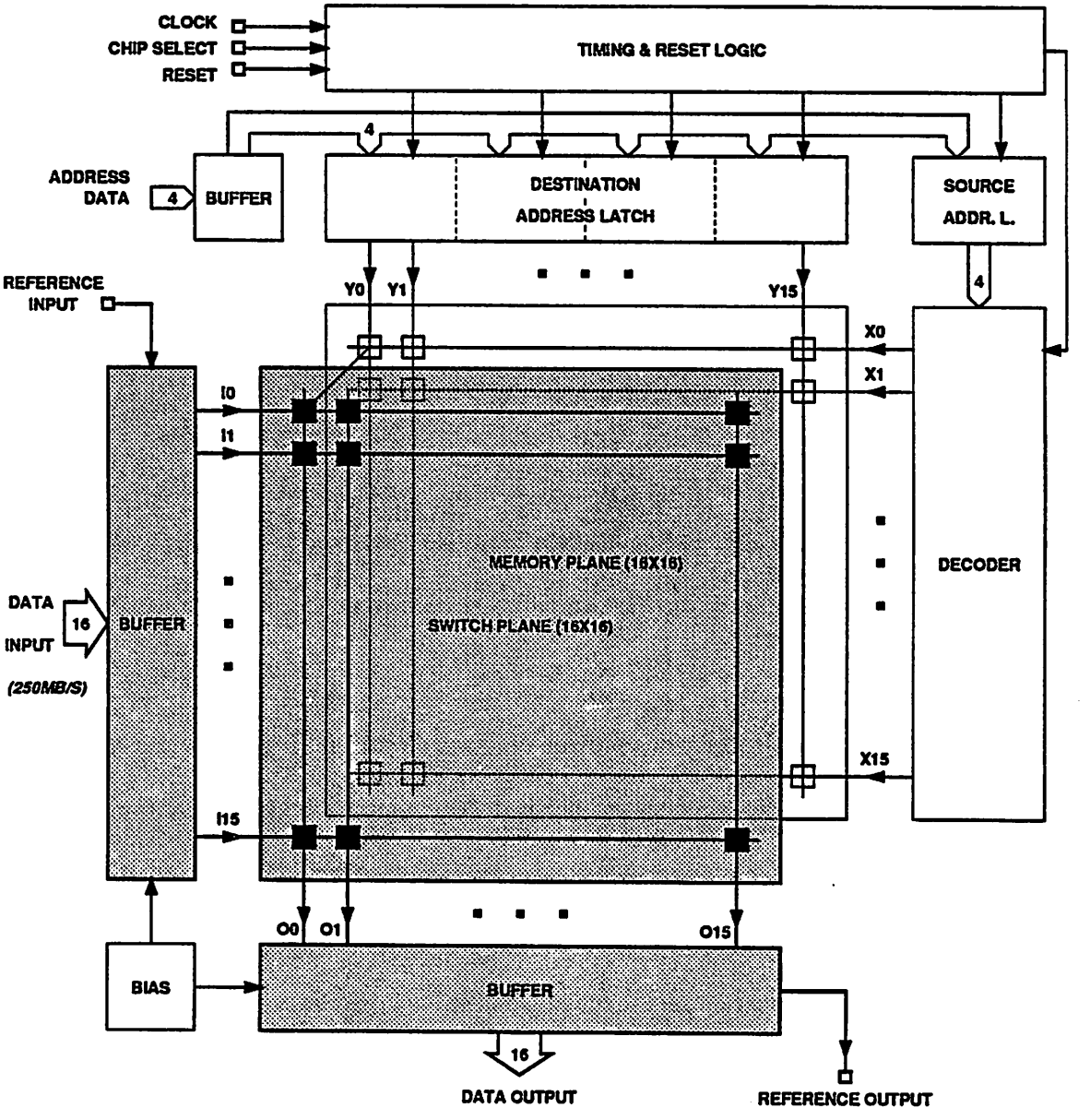


Figure 5.1 Prototype chip architecture where shaded area represents high-speed data paths.

The number of stations that this single switch connects is limited to 16. But because this chip is designed to support the expansion mechanisms illustrated in Figures 2.11 and 2.12, switching networks of sizes greater than 16 can be built. Although fewer chips are required to construct a large switching network if the number is increased, the total pin count increases

rapidly and a larger package is needed. For example, if the switch size is increased from 16×16 to 32×32, the pin count increases from 48 to 80. One disadvantage of a larger package is the higher lead inductances associated with it.

This architecture is suited for the access method shown in Figure 2.7, in that any hardware for access request is not included. Access requests should be handled in a host controller, and the resulting switch setup information must be provided through the clocked control bus. The communication rate of the synchronous control bus is only 10 MHz. Compared to this, data on 16 high-speed channels are asynchronous.

5.1.1 High-Speed Paths

The high-speed paths include the crosspoints organized into a matrix, the input buffer that drives the crosspoints, and the output buffer that amplifies signals from the matrix to drive external loads.

In this design, the high-speed paths provide only waveform amplification and level restoration for the passing data. No circuitry is included for retiming or synchronization. This is because, at the desired speed, timing recovery and regeneration of the incoming or outgoing data, i.e., retiming function, would require more technology bandwidth. Also, because the data on 16 channels are asynchronous, the retiming function would be required in every channel. Since timing recovery at this high speed is costly, it cannot be implemented entirely on the chip. In addition, this function is not needed for every switch chip if timing jitter addition from each chip is minimal. It may be economical to provide the function at each station or at nodes where the accumulated jitter and noise are excessive, and to minimize the jitter and noise in individual chip. In this switch, the maximum jitter is specified as 10% of the bit period of data, which is adequate for proper timing recovery.

The input and output buffers are designed to be voltage level compatible for easy expan-

sion. Also even though levels do not match, their voltage swing is made compatible with the ECL swing; that is, the input accepts ECL swing and the output can drive ECL, if properly level shifted. For this purpose, a differential stage is used at each chip input. To reduce the number of pads, however, one terminal of the differential inputs is connected to the common node REFERENCE INPUT. A reference voltage that is the center of the input swing is normally provided to this node. In cases of cascading the switch chips, the reference from this chip (REFERENCE OUTPUT) that nominally represents the center level of the output swing can be connected to the REFERENCE INPUT node of the next chip. Combined with this reference output, the output may drive any differential input stages such as line receivers and LED or laser diode drivers directly.

5.1.2 Memory and Control Blocks

The memory and control blocks include the memory matrix that stores switching information and the latches, decoder, and timing and reset logic that are needed to update the information. In this chip, the memory matrix consists of static storage cells that keep information as long as power is supplied. To update the information stored in the matrix or set up the switch, a new information must be clocked into the source and destination address latches from the 4-bit ADDRESS DATA bus and, in turn, written into the memory. The timing and reset logic generates a sequence of pulses needed for this task from the external clock in combination with other control signals such as CHIP SELECT and RESET. Because setup requests in this architecture are initiated by source stations, the source address latched needs decoding while the destination addresses do not.

Basically, the memory and control blocks operate at low speed. The worst case maximum clock rate is specified to be 10 MHz to allow good interfaces to microprocessor busses or other computer communication channels. The inputs to these blocks are TTL or CMOS compatible.

5.1.3 Setup Operation

There are two setup modes for this switch: quasi-serial and parallel. Because the crosspoint cell is bit-mapped to the memory cell, to set up a path, the corresponding memory cell is written with '1'.

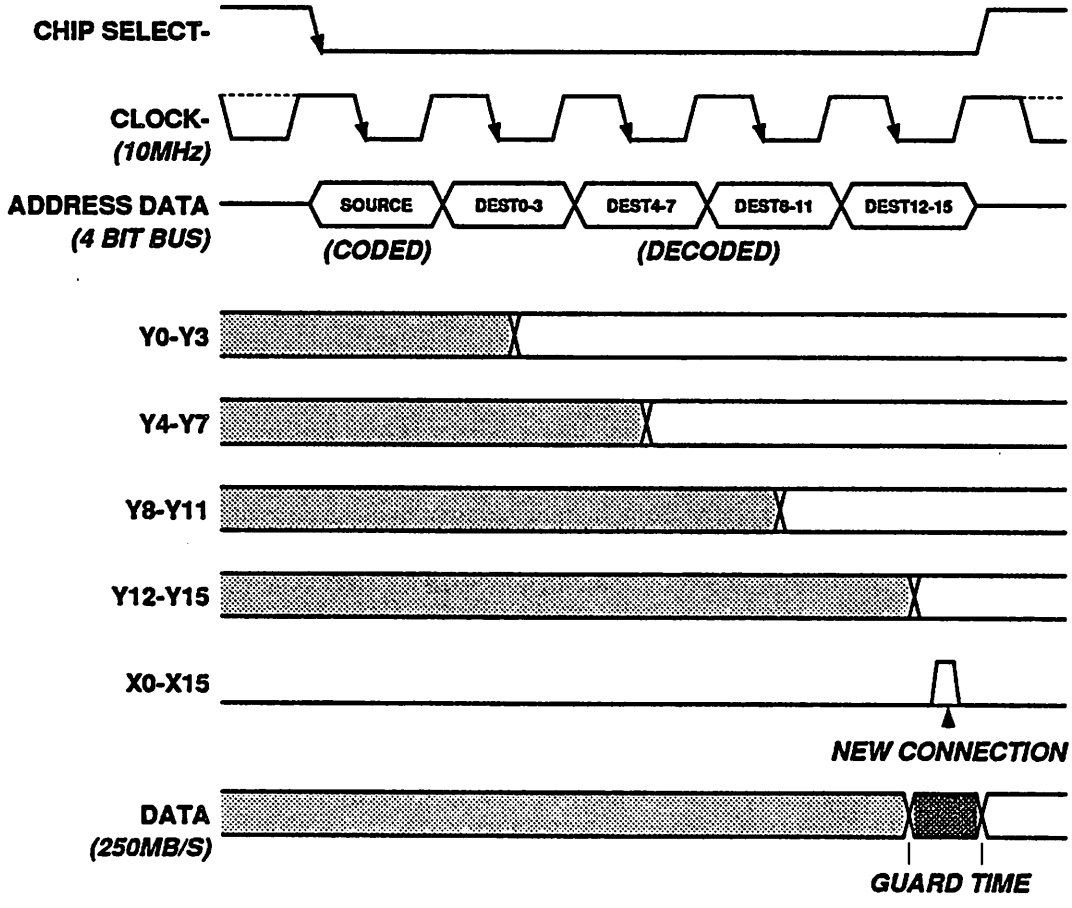


Figure 5.2 Quasi-serial setup mode needs 5 clock cycles to establish a new connection.

In the quasi-serial mode, any setup including selective or full broadcasting needs 5 clock cycles or 20 bits of information, first 4 bits specify a coded source address and the rest, decoded destination addresses. The source address is latched into SOURCE ADDRESS LATCH by a pulse derived from the first clock edge. While it is being decoded, bits 0-3 of destination addresses get into the corresponding positions in DESTINATION ADDRESS LATCH by

another pulse generated from the next clock edge. This repeats until bits 12-15 of destination addresses are loaded. Then, a final pulse activates the decoded source address line (one of X0 - X15), which in turn writes the 16-bit wide, latched destination addresses (Y0 - Y15) into the row of memory. This concept is illustrated in Figure 5.2. The 16-bit wide destination addresses can be of any bit patterns. If they are all 1's, the source is broadcast to all destinations.

Since the specified clock frequency is 10 MHz, one setup takes 500 ns in this setup mode. As a result, if the high-speed data are packetized, the packet size should be large enough so that the setup time is a negligible portion of the total transmission time. In fact, guard time should be included in one transmission. During the guard time (or dead time) when a new connection is made, data cannot be sent reliably. A good value for 200 Mb/s transmission would be 1 kbits.

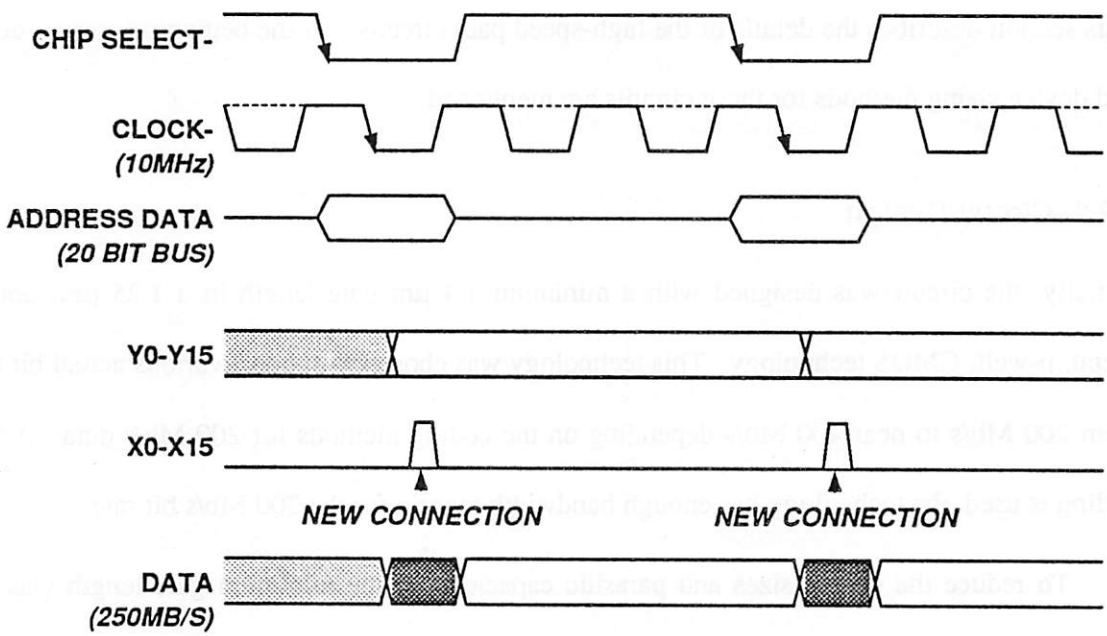


Figure 5.3 Parallel setup mode makes a connection in one clock cycle.

The parallel mode allows a setup in one clock cycle as depicted in Figure 5.3. The 4 bit source address and the 16 bit destination addresses are latched into at the same time. This mode is added for debugging purposes and would be useful if the timing logic fails to produce the

sequence of pulses. Due to the bonding pad limitation on the experimental chips, however, the parallel destination addresses are loaded through internal probe pads. If this setup mode is used, the minimum packet size can be reduced by a factor of 5.

Clearing of entire memory or disconnecting all crosspoints is done by activating the RESET pin. Also, for convenient checking of every link between the station and the switch, LOOP BACK feature is included. In this mode, diagonal cells in the memory matrix are written with '1' at once. As a result, I0 is switched to O0, I1 to O1, I2 to O2, and so on; i.e., the data sent from one station return to itself if the loop is not broken.

5.2 HIGH-SPEED PATH CIRCUITS

This section describes the details of the high-speed path circuits. At the beginning, circuit design and device sizing methods for these circuits are mentioned.

5.2.1 Circuit Design

Initially, the circuit was designed with a minimum 1.4 μm gate length in a 1.25 μm, double-metal, p-well, CMOS technology. This technology was chosen to support various actual bit rates from 200 Mb/s to near 400 Mb/s depending on the coding methods for 200 Mb/s data. If NRZ coding is used, the technology has enough bandwidth margin for the 200 Mb/s bit rate.

To reduce the device sizes and parasitic capacitances, the minimum gate length was normally utilized. However, because the minimum length devices have the short channel effects [47], the less accurate, conventional models (level 1 and level 2) of SPICE2 [52] have been rarely used for hand calculation or simulation. Rather, the design has been based on reference device characteristics and extensive circuit simulations carried out using SPICE2 with BSIM [53] parameters. The BSIM parameters have been extracted from typical test wafers provided by the fabrication site. The I_{DS} vs. V_{DS} curves for 1.4 μm channel length MOSFETs are shown in Figure

5.4.

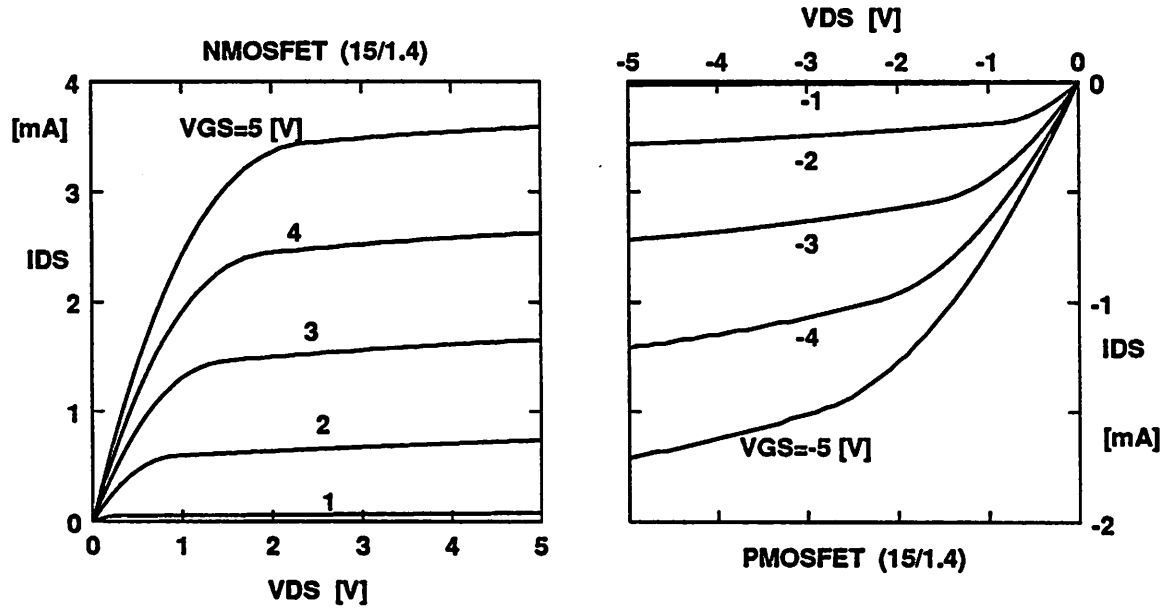


Figure 5.4 Drain current vs. drain voltage curves for reference MOSFETs.

Primary design requirements for the high-speed path circuits are maximum bandwidth and symmetric rise/fall characteristics. Here, symmetric rise/fall is required to maintain the duty cycle of the signal and maximize the bit rate of the switch. The bandwidth is maximized by using minimum channel length devices and maximum possible gate biases. And symmetry is obtained by properly ratioing p- and n-MOSFETs so that the logic threshold is close to $\frac{1}{2} V_{DD}$. If so for an ordinary CMOS inverter, the bandwidth is maximized because this logic threshold guarantees maximum biases for both transistors.

In this technology, a 1.4 μm gate length inverter with a p-MOSFET to n-MOSFET width ratio of 2.5 has a logic threshold of ~2.5 V (see Figure 4.25). This inverter is a reference for the design of transresistance amplifiers and matching inverters in the high-speed paths. As will be clear in later subsections, for this prototype, the transresistance amplifier shown in Figure 4.27 is applied extensively. Resistors in the transresistance amplifiers are also designed from a reference

device and trimmed via simulations.

5.2.2 Switch and Sense Amplifier

As discussed in Section 4.1, complexity and area of the switch cell should be minimized for maximum speed and reduced parasitics. The 4-transistor tri-state inverter shown in Figure 4.3 (b) is selected for the switch cell, because this switch cell has wide bandwidth, small area, low parasitic capacitances, low power consumption, and negligible feedthrough, as pointed out in Section 4.1.2. Also, unlike the passive switch cells, this does not cause loading fluctuation on the input line.

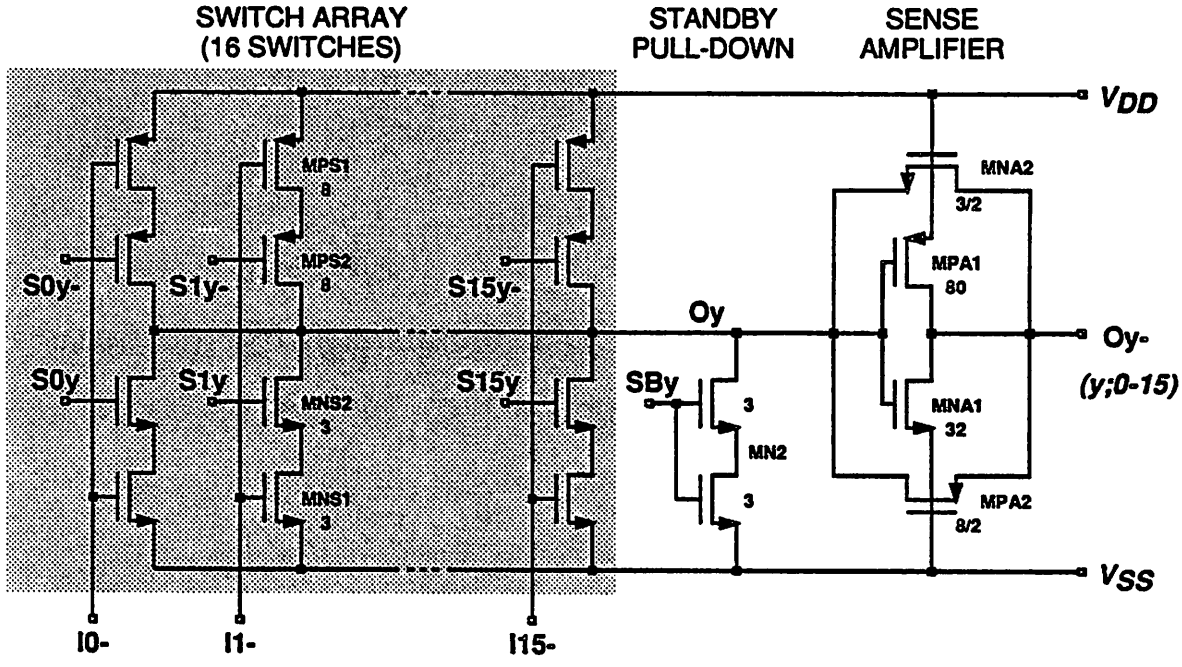


Figure 5.5 Switch and sense amplifier circuits.

Figure 5.5 shows the switch cells in the shaded area and the sense amplifier circuits on the right. It indicates that at each output line of the crosspoint matrix, a CMOS transresistance amplifier senses current from one of the 16 switch cells. The transistors MNS2 and MPS2 are switches controlled by the stored value Sxy and its complement Sxy-, respectively. If they are

ON, the high-speed input signal Ix- is inverted at the output Oy.

The transistor MN2 is used to pull Oy down during a standby state, when none of the 16 switches connected to this output is ON. If MN2 is not present, current into the transresistance amplifier is zero and Oy- sits at the logic threshold voltage. This undefined state in digital circuits is prevented by a NOR gate and MN2, as shown in Figure 5.6. The NOR gate is realized with a pseudo-NMOS technique to reduce routing complexity in the matrix. MN2 is implemented as two minimum size transistors in series, to match the sink current level.

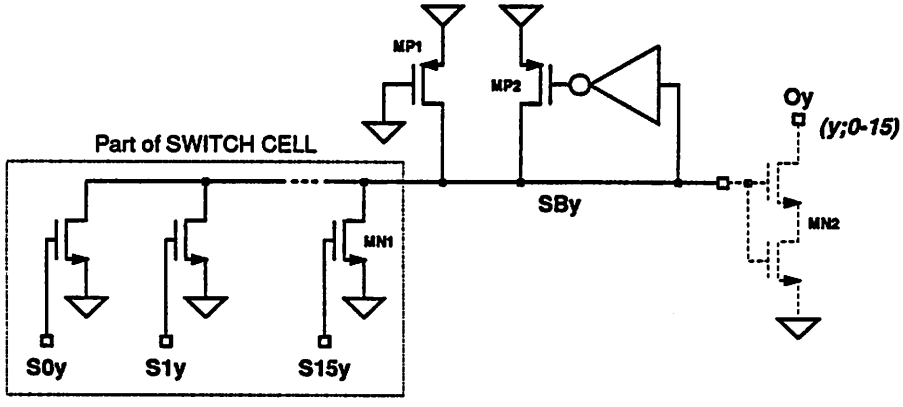


Figure 5.6 Standby pull-down circuit.

The design starts by approximating parasitic capacitance and estimating voltage swing at Oy. The capacitance is mainly contributed from drain junctions of the switch transistors, output routing, and gates of the amplifier. But since the transistor sizes are unknown, it is hard to estimate their contribution. Instead, the total capacitance is roughly approximated based on the routing capacitance. If the first metal is used for output routing, the capacitance for a 2 mm long, minimum width wire is about 100 fF. Hence, the total capacitance would be less than 1 pF. The voltage swing depends on a desired swing at Oy-. Remember that the inverter (MNA1 and MPA1 pair) used as an amplifier has a gain about 10 for a limited range of output voltage, i.e. approximately from 1 to 4 V. If this swing is desired and gain degradation due to feedback load-

ing is minor, the swing at Oy is on the order of 0.5 V.

Current from the switch to drive Oy in 1.5 ns is then about 330 μA, which allows the use of minimum size (3/1.4) for both MNS1 and MNS2. MPS1 and MPS2 are sized as 8/1.4 for symmetric current drive. The feedback resistance for a 3 V swing at Oy- is about 5 kohm. Final device sizes after fine tuning are indicated in Figure 5.5. The gate lengths are not shown unless they deviate from 1.4 μm.

One thing to note is that the feedback resistance changes with process and power variations. However, as discussed in Section 4.4.7 in relation with Eq. (4.25), the variations affect the crosspoint cell current similarly, resulting in a less sensitive voltage swing at Oy-.

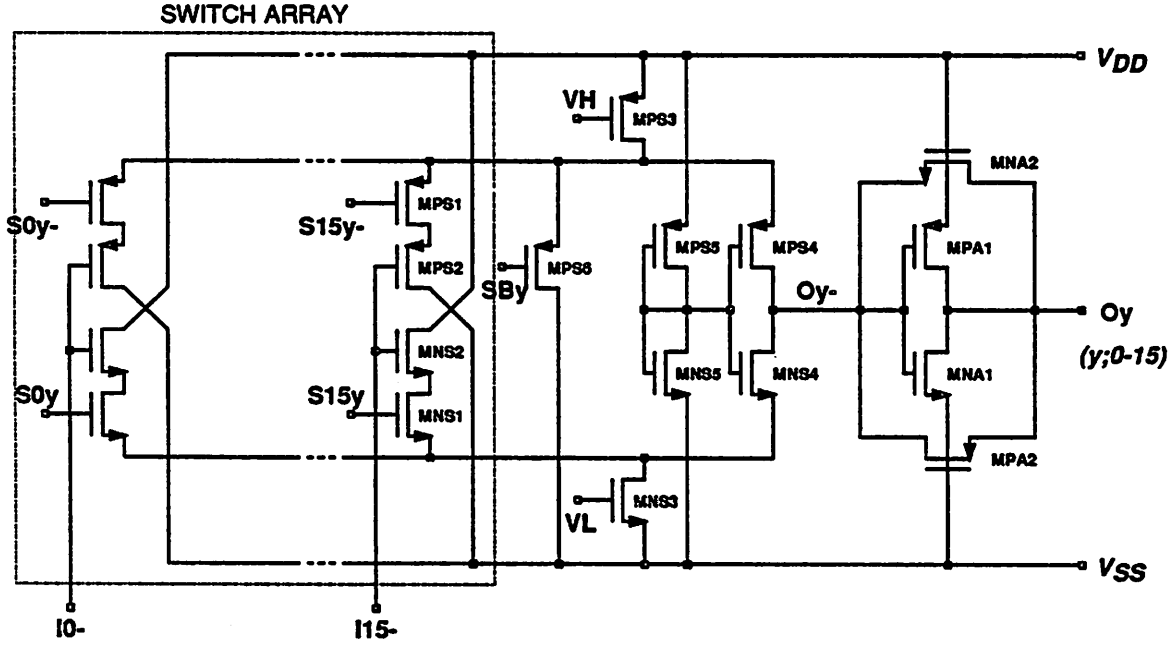


Figure 5.7 Constant current steering switch and sense amplifier.

Although the switch does not steer constant current, the supply current transient is very small due to the small charging current and the voltage clamping at Oy. If constant-current steering is desired, it can be done by using a modified switch cell depicted in Figure 5.7. Here, VH and VL are biases for the current sources MNS3 and MPS3. And the MNS5 - MPS5 pair

provides a reference equal to the logic threshold for MNS4 and MPS4. MPS6 is used for standby pull-down. This is basically the current steering circuit in Figure 4.18 where MN+ and MP+ are replaced by 16 switch cells. Major disadvantages of this circuit include a lower speed arising from the increase in complexity and device sizes. Routing is also more complicated. Because these factors more than offset the benefit gained, this circuit is not adopted for the prototype.

5.2.3 Input Buffer

For proper interfacing between the two identical switch chips as needed for construction of a larger switching network, the input buffer is designed to be level compatible with the output buffer. In this prototype, the output has a nominal swing greater than, but close to 1 V, which is centered at the logic threshold of the reference inverter. Thus, the input buffer should discern high/low levels less than 1 V apart. By making the input swing ECL compatible, this prototype may be interfaced to ECLs with proper DC level shifting.

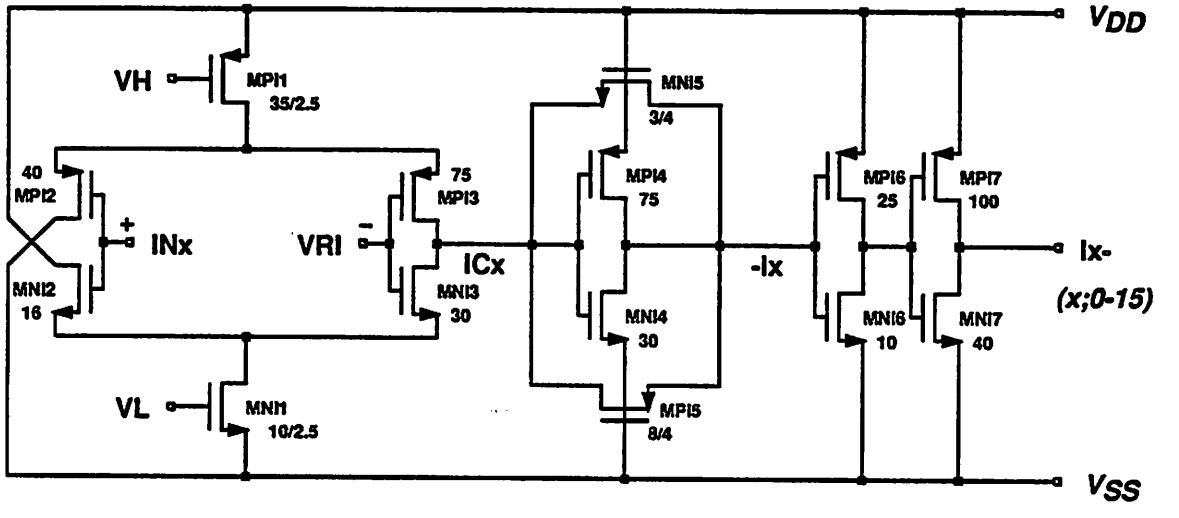


Figure 5.8 Input buffer circuit.

Figure 5.8 shows the input buffer composed of the constant current steering circuit in Figure 4.17, the transresistance amplifier in Figure 4.27, and two inverters. The differential input stage has been selected to facilitate the small input swing. However, to reduce the pin count, the negative terminal of the differential input stage is connected to a common input reference VRI. MNI1 and MPI1 are mirrored current sources. The transresistance amplifier connected to the output of the current steering circuit serves rather as a current-to-voltage converter than as a voltage swing limiter for ICx. That is, it transforms the bidirectional current into a bi-level voltage. In fact, the capacitance at ICx is not large enough to need swing reduction. The converted voltage -Ix drives a pair of inverters that provide full 5 V swing on Ix-. If the swing at -Ix is large, static power consumption in the first inverter decreases. In this particular design, the switched current level is about 160 μA, the feedback resistance is 8 kohm, and the voltage at -Ix swings between 1.3 and 3.6 V.

Input voltage difference needed to steer current I_n fully from MNI2 to MNI3 is

$$\Delta V_{INx} = \left[\frac{I_n}{\frac{k_n'}{2}(\frac{W}{L})_2}\right]^{1/2} + \left[\frac{I_n}{\frac{k_n'}{2}(\frac{W}{L})_3}\right]^{1/2}. \qquad (5.1)$$

Therefore, the MOSFETs for current steering must be made wider to lower the required input swing. However if this swing is too small, since the widths increase by the square of required swing reduction, the gate, source, and drain capacitances increase rapidly and the bandwidth of this circuit decreases. Unlike a case when both input terminals move opposite and symmetric way, even the capacitances at the common source nodes (sources of MNI2 - MNI3 pair and MPI2 - MPI3 pair) hurt the bandwidth. These nodes need to follow the input INx quickly for fast steering.

This also puts a limit on how small the saturation voltages of MNI1 and MPI1 and how large these MOSFETs can be. The saturation voltages may not be set too high, though, to prevent the transistors from getting out of the saturation. Current sources in the linear region have

characteristics far from ideal. Although better current source characteristics can be obtained by using longer channel devices, corresponding width increase is not desirable.

In this prototype, saturation voltages of the current sources are about 1 V, and ΔV_{INx} is about 0.9 V. The MOSFET sizes are shown in Figure 5.8. Note that the MOSFETs on the input side (MNI2 and MPI2) are sized smaller than MNI3 and MPI3 to reduce the input loading.

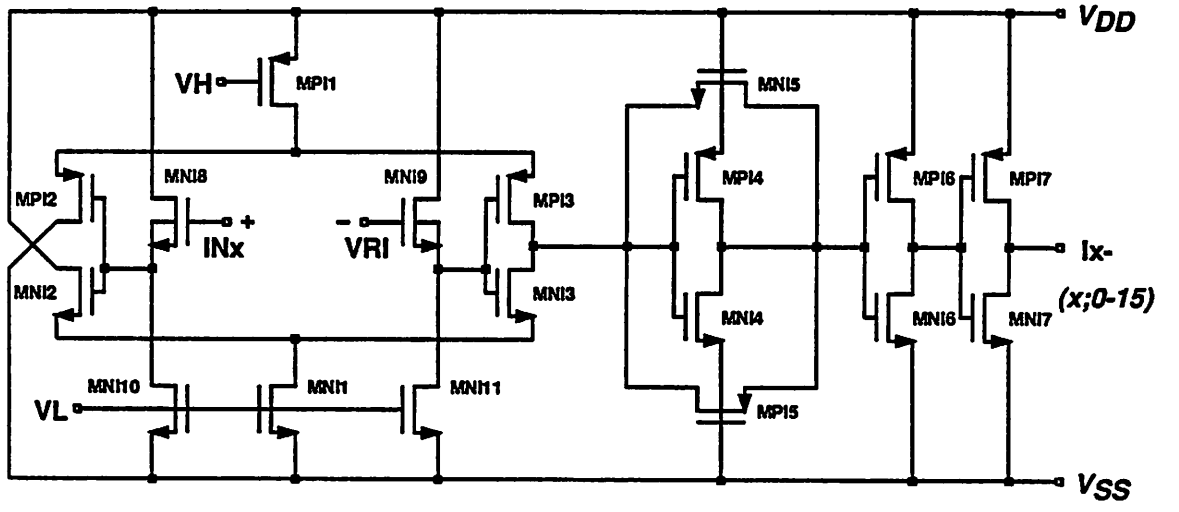


Figure 5.9 ECL compatible input buffer.

Extension of this circuit to an ECL compatible one can be done easily by adding two source followers at the front end, as shown in Figure 5.9. High gain n-MOSFETs with the wells tied to the sources (MNI8 and MNI9) facilitate good followers.

5.2.4 Bias Circuit

The bias voltages VH and VL are generated in the circuit shown in Figure 5.10. Note that the current in MNB1 is equal to the one in MPB1. Also note that MNB1 and MNI1 in the input buffer form a current mirror, like MPB1 and MPI1 do. Hence, any current variation in MNB1 affects I_n and I_p equally.

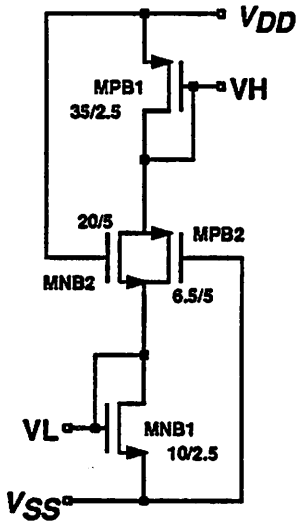


Figure 5.10 Bias generating circuit.

The MNB2 - MPB2 pair is a CMOS transmission gate resistor whose value is determined by Eq. (4.23). As a result, the resistance varies depending on the supply voltage. If VH and VL are not sensitive to the supply variation, the current through this resistor changes inversely with the supply. This current is mirrored to the current steering circuit of the input buffer, and converted to a voltage swing by the transresistance amplifier. As discussed in Section 4.4.7 with relation to Eq. (4.25), the sensitivity of this voltage swing to the supply variation is equal to that of VH-VL.

If MNB1 and MPB1 are hard limiting diodes, any change in $V_{DD} - V_{SS}$ will appear equally in VH-VL. Since VH-VL is $(V_{DD} - V_{SS} - V_{GSnb1} - |V_{GSpb1}|)$, the sensitivity of VH-VL is greater than 1. However, the transistors are not hard limiting in practice. For example, if the supply voltage drops, VH-VL tends to drop. But this reduces current in the MOSFETs, which in turn reduces the gate-source voltages. Thus, the drop of VH-VL is smaller than that of the supply. Furthermore, as the supply drop increases resistance of the MNB2-MPB2 pair, the current and V_{GS}'s are reduced more. This suggests that the sensitivity is lower than the hard limiting case. To a first order, if MNB1 and MPB1 are made weaker by decreasing the aspect ratios, the

sensitivity falls far below 1.

The use of weak transistors is limited, however, by the increase in the saturation voltages. If they are too high, dynamic range of the input circuit will be restricted. On the other hand, a moderately high saturation voltage helps to reduce the size of the current source. In this design, VH and VL are set about 3 and 2 V, respectively.

5.2.5 Output Buffer

The output buffer has the same structure as the input buffer except the last two inverters removed, as shown in Figure 5.11. But main purpose of this combination is to generate the output swing levels approximately +0.5 V (high) or -0.5 V (low) beyond the logic threshold voltage of the reference inverter. For a symmetric design, these would be close to 3V and 2V. Remind that, to a first order, the swing is not sensitive to process and supply variations due to the bias scheme explained in the previous section, which controls current in the steering circuit to be inversely proportional to the feedback resistance. However, due to a large external capacitance, the output may show ringing (see Figure 4.29). Because the chip provides the logic threshold as an output reference, the outputs can readily be interfaced to the inputs of other identical chips or LED/laser diode drivers with differential inputs.

One of the input nodes of the steering circuit is connected to the internal reference identical to the logic threshold of the MNO6 - MPO6 pair. This reference is stabilized by adding the bypass capacitors C_s and C_d. To reduce the capacitance at Oy- that is driven by the sense amplifier (Figure 5.5), MNO2 and MPO2 are sized much smaller than MNO3 and MPO3. Also, to minimize capacitances at the common source nodes, minimum gate length devices are used for the current sources. Although the current mirroring becomes non-ideal, this design increases the bandwidth.

The small swing at the output enables us to achieve the speed introducing only 1/5 of tran-

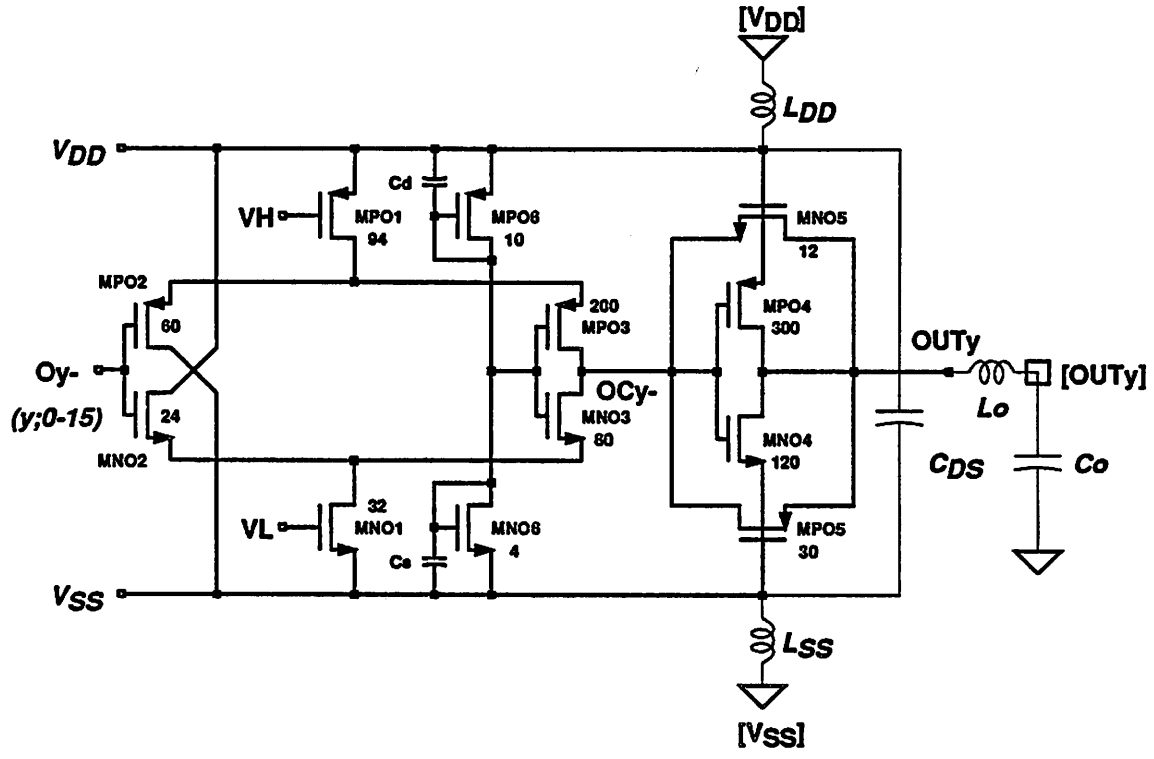


Figure 5.11 Output buffer circuit.

sient current and power line noise compared to the conventional 5V swing. As well, it reduces dynamic power consumption at the output circuitry by a factor of 25. For a 200 MHz clock and a 10 pF load, the dynamic power per output is only 2 mW. However, since both n- and p-channel transistors of the last inverter are ON all the time, large static power is consumed instead. Current level in the MNO4 - MPO4 pair is about 8.8 mA, and the steered current is near 1 mA. Most of the chip power is dissipated in this circuit.

The resonance of the output parasitics L_o and C_o may complicate the output waveform. For L_o = 5 nH and C_o = 10 pF, the resonance frequency is close to 700 MHz. Since this is about 3.5 times higher than the maximum bit rate, the effect would be minor.

Roughly approximated ECL levels can be obtained using the output circuit in Figure 5.12. The diode D may be realized using a well-to-substrate junction. This circuit is not a perfect solu-

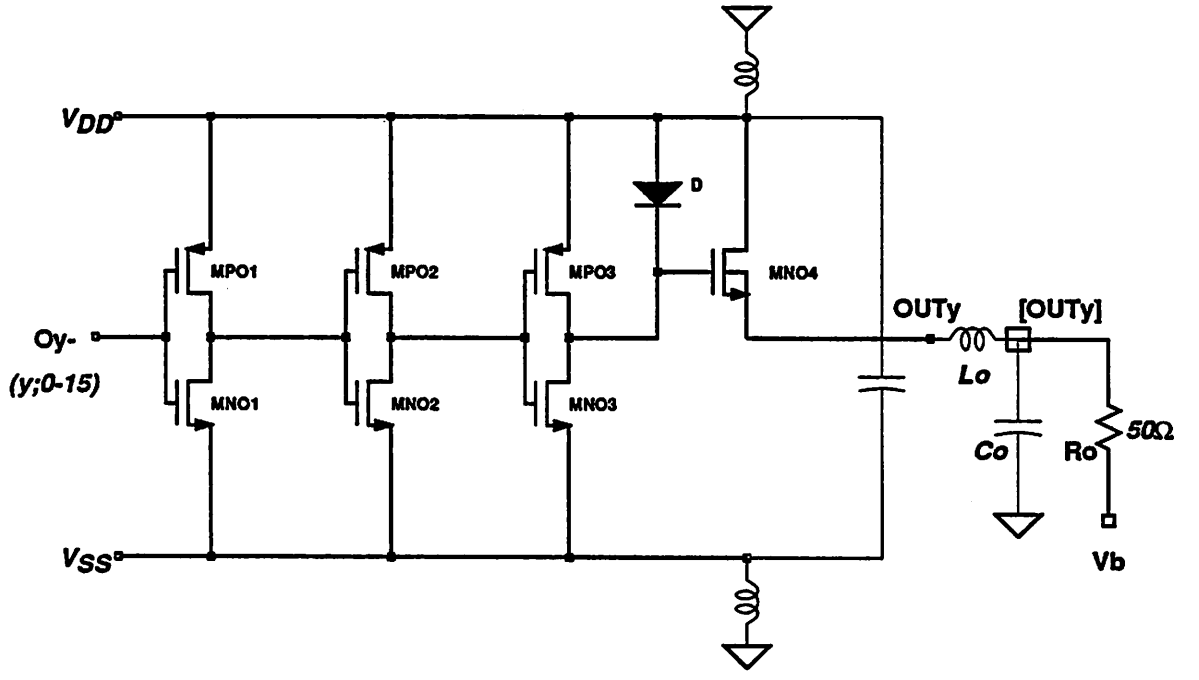


Figure 5.12 Output buffer with approximate ECL levels.

tion because the output levels are sensitive to V_{Tn} variation of MNO4. If the high level is too high, it may saturate the input transistors of an ECL gate and slow down the speed.

5.2.6 Reference Output

The reference output circuit is equal to the last stage of the output buffer, i.e. MNO4 - MPO4, except that its input and output nodes are shorted together. Automatically, this circuit generates the logic threshold of the inverter as the reference. Hence from Appendix 2, the reference voltage is represented as

$$V_{RO} = V_{LT} = \frac{\beta_R^{1/m} V_{DD} + V_{Tn} - \beta_R^{1/m} |V_{Tp}|}{1 + \beta_R^{1/m}} \tag{5.2}$$

where m = 1 or 2 depending on the channel length and

$$\beta_R = \frac{k_p' (\frac{W}{L})_p}{k_n' (\frac{W}{L})_n}. \tag{5.3}$$

Note that if $\beta_R = 1$, $V_{RO} = \frac{(V_{DD} + V_{Tn} - |V_{Tp}|)}{2}$.

5.2.7 Simulation Results

The high-speed path circuits have been combined and simulated using SPICE2 and the BSIM parameters. In the simulation, 200 Mb/s RZ data with maximum transition density (or 400 Mb/s NRZ) has been used. The parasitics assumed are $L_{DD} = L_{SS} = 2$ nH, $L_o = 5$ nH, $C_{DS} = 160$ pF, and $C_o = 10$ pF. Figure 5.13 shows the simulation results. For the best case when only one channel is active, the output is under damped and power supply noise is negligible (about 22 mV). The propagation delay through this switch is about 5.2 ns.

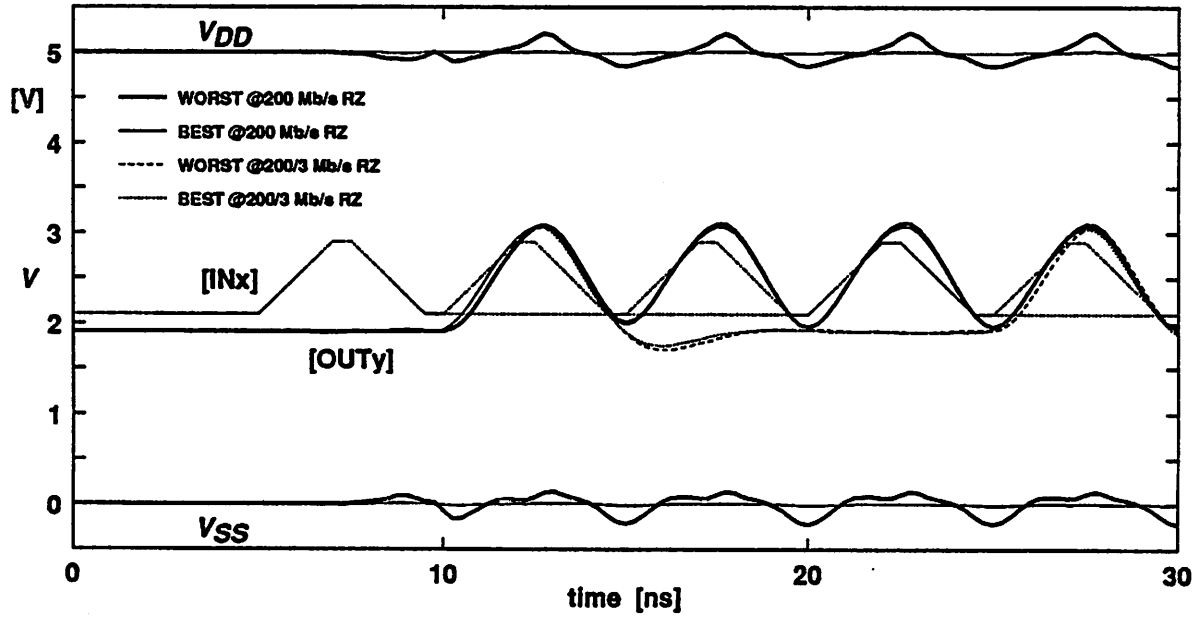


Figure 5.13 Simulation results for the high-speed path circuits.

To see the timing jitter arising from intersymbol interference, the input was changed so that it has one pulse in every 3 bit-periods. The intersymbol interference mainly comes from the under damping characteristics of the high-speed paths (actually output stage). If the bit rate is 400 Mb/s, the circuit cannot settle before the next pulse arrives, and ringing on the signal does

not decay in one bit-period. However, for a bit rate sufficiently low, the circuit settles within one bit-period and past history of the pulses dies out. From the curves, the jitter is measured as less than 1 %.

Worst case simulations have been carried out by assuming that one input is broadcast to all 16 outputs. The power supply inductances effectively increase by 16 and the bypass capacitance decreases by 16 in this case. Although the power supply noise is increased to 0.38 V p-p, the output waveform is very close to that of the best case. Crosstalk through power supply fluctuation has been simulated by imposing high-frequency, 0.5 V p-p sinusoidal signals on the power lines. The frequency is either 400 MHz or 444 MHz, and the phase difference between V_{DD} and V_{SS} is 0 or 180 degrees. Maximum jitter from various combinations of the frequency and phase is still less than 5 %.

5.3 MEMORY AND CONTROL CIRCUITS

The memory and control circuits are basically conventional CMOS logic circuits. Therefore, only unique features of this prototype will be described. For these circuits, the p-MOSFET to n-MOSFET aspect ratio of 2 was maintained mostly to reduce the area.

5.3.1 Memory Cell

As discussed in Section 4.2, complexity and area of the memory cell should be minimized to facilitate routing of required signals and reduce parasitics. This is especially enforcing, because the crosspoint switch and memory are merged into a cell that is used to construct a 2-dimensional matrix.

Among many choices of memory circuits, the modified static latch in Figure 4.5 is advantageous because of the reasons mentioned in Section 4.2.2: six small (mostly minimum) devices, a word line, a bit line, no clock skew problem, and easy design. Although it has drawbacks such as

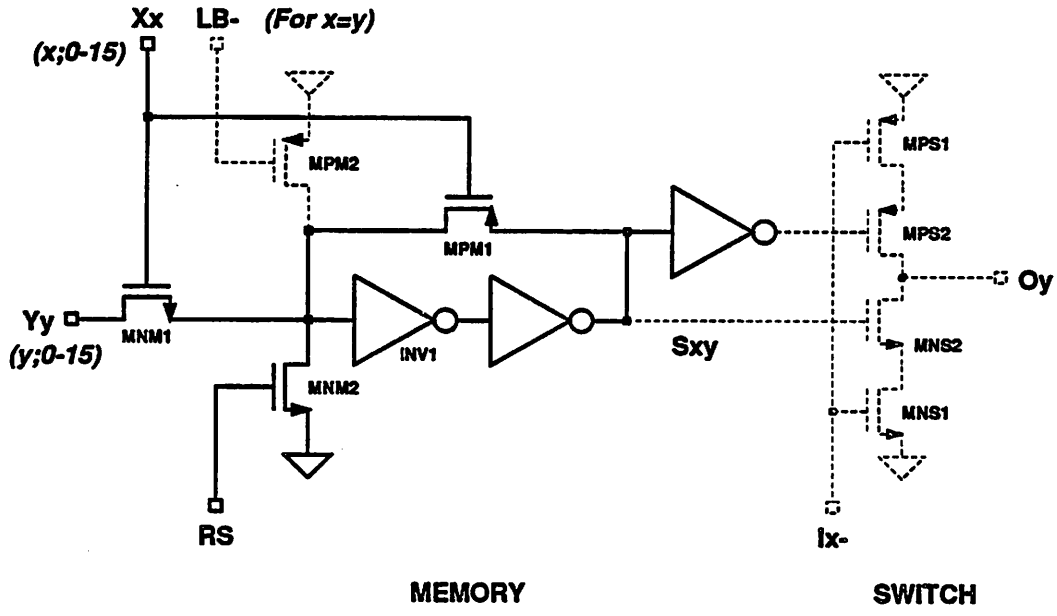


Figure 5.14 Memory cell.

reduced noise margin and potential static power dissipation, their effect on memory operation and already large total power consumption would be minor. Most of the chip power is dissipated in the high-speed paths.

Figure 5.14 shows the memory cell combined with the switch cell. In this memory, the fully decoded and pulsed source address Xx either gates the destination address information Yy through MNM1 or latches the circuit via MPM1. To reset the entire array at once, MNM2 is provided. Also, for cells at diagonal locations, i.e. $x = y$, MPM2 is added to set up a loop back mode.

Except INV1, MNM2, and MPM2, minimum size devices are used in this circuit. INV1 needs a low logic threshold, because initial high level at its input is $V_{DD} - V_{Tn}$. However, the logic threshold may not be too low, since the low level at its input may drift up to $V_{SS} + |V_{Tp}|$. A p- to n-MOSFET ratio of 1.5 is proper for this inverter. MNM2 and MPM2 need to be large to override already stored information.

5.3.2 Destination Address Latch

Destination address latch stores the destination selection information Yy and has two input ports: ADi for the quasi-serial loading and ADy for the parallel loading. As shown in Figure 5.15, this latch is a modified version of the memory cell. Note that the two gating n-MOSFETs are connected in parallel and the latching p-MOSFETs, in series. The latch is reset at the same time when the memory matrix is cleared.

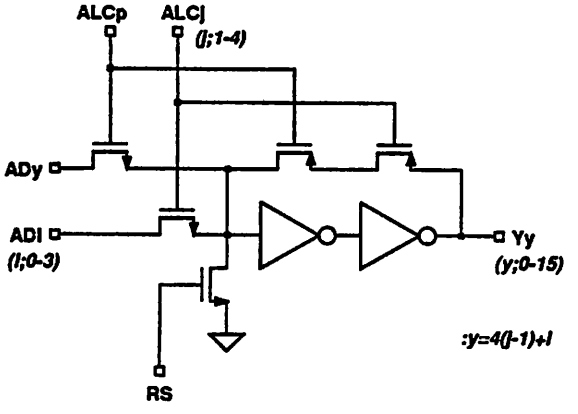


Figure 5.15 Destination address latch cell.

In the quasi-serial mode, the pulse ALCj gates AD0 - AD3 into 4 latches. ALC1 is activated first to update Y0 - Y3, ALC2 next for Y4 - Y7, and so on (see Figure 5.2). Whereas in the parallel mode, ALCp loads all 16 latches at once as explained in Figure 5.3.

5.3.3 Source Address Latch and Decoder

Source address latch and decoder circuits generate the pulse Xx (Figure 5.16). The latch is basically identical to the destination address latch, except that its inputs are tied together and reset function is removed. ALC0 or ALCp is used to clock in ADi depending on the setup mode. The 1 of 16 decoder following the latch is a fully static CMOS circuit. And the write pulse WP- gates the output of this decoder.

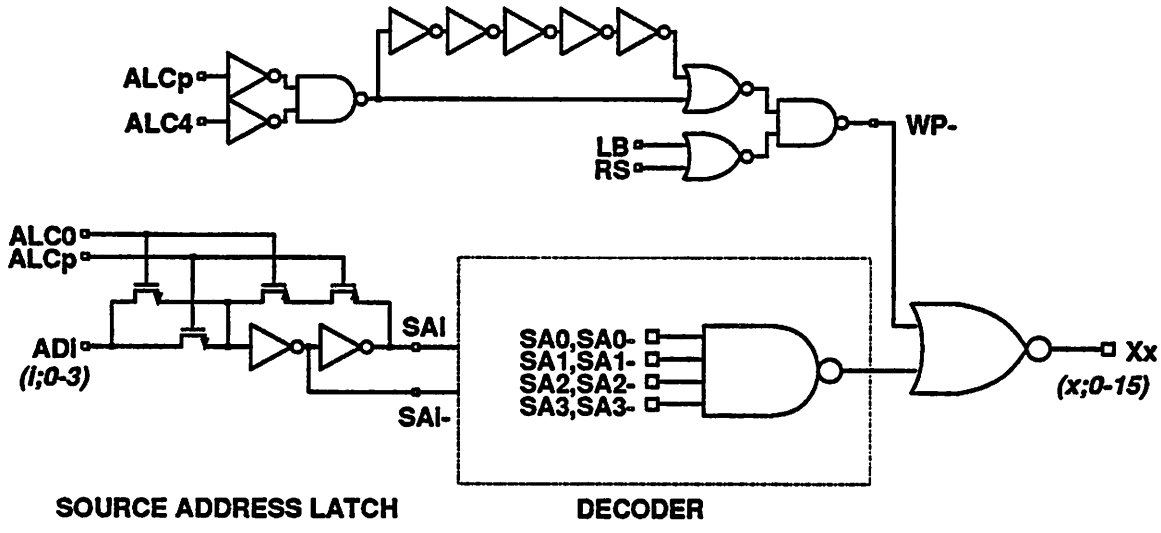


Figure 5.16 Source address latch and decoder circuits.

WP- is generated from the falling edges of ALC4 or ALCp, and the pulse duration is determined by the 5-stage inverter delay.

5.3.4 Timing and Reset Control Logic

The control signals for latching and resetting in the above circuits are derived from the logic shown in Figure 5.17. A 5-bit serial shift register is used to generate the sequential pulses ALC0 - 4 from the clock ([ADS-]) edges. The chip select signal ([CS-]) enables the clock when it is low. Also this signal or the reset ([RS-]) clears the shift register for the start of a new sequence. The loop back mode is controlled by [LB-], and the parallel/ quasi-serial loading is selected by [P/QS-].

The master-slave D flip-flop used in the shift register is realized by cascading two latches (one with p-MOSFET gating and the other with n-MOSFET gating) as shown in Figure 5.17. This is simple and free from the clock skew related race problem because it needs only one clock line.

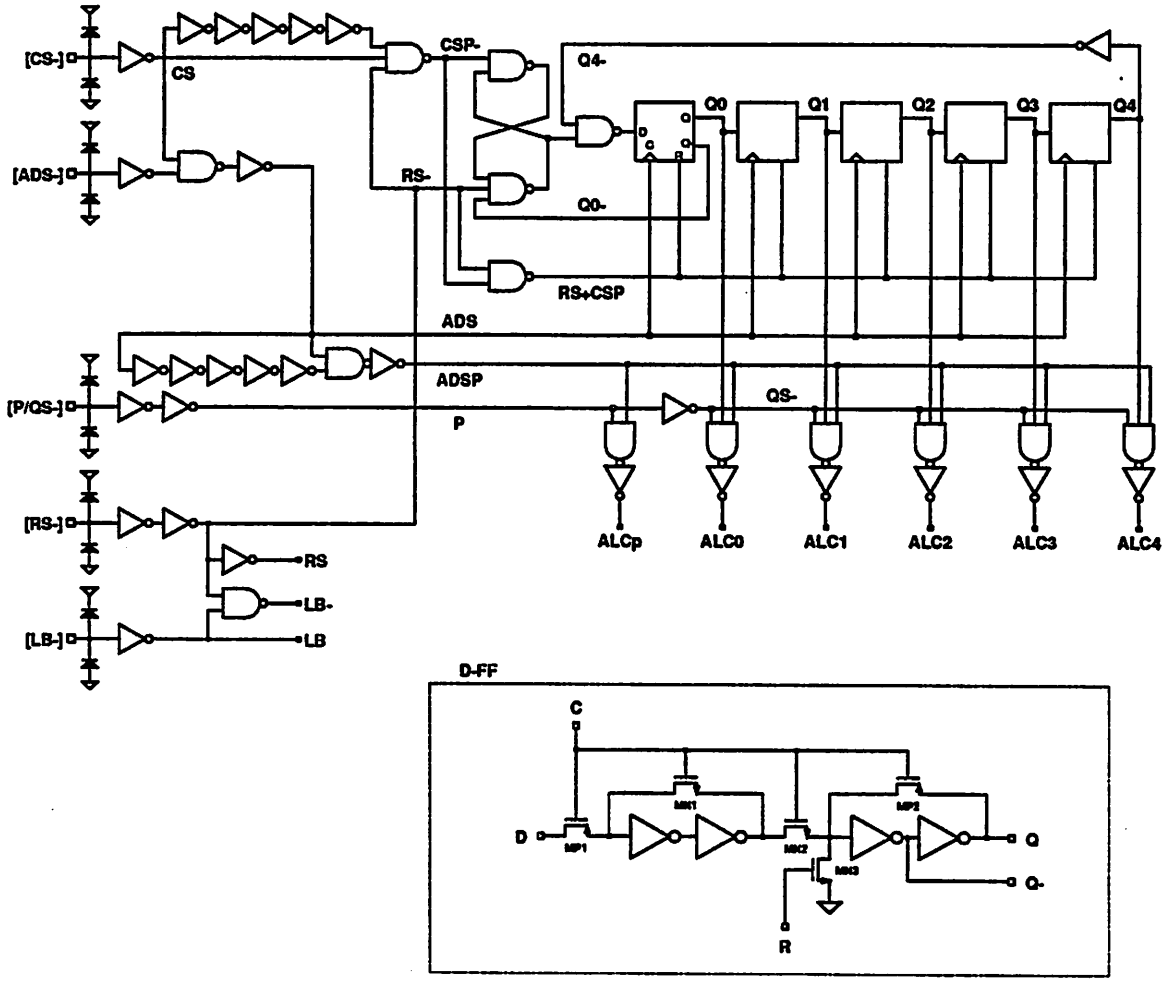


Figure 5.17 Timing and reset control logic.

Since all pulses are generated from the transition edges and their widths are determined by the inverter delays, the speed of this control logic is approximately proportional to the technology bandwidth. In this prototype, the clock frequency can go up to 50 MHz based on the simulation.

5.4 LAYOUT AND FABRICATION

5.4.1 Layout

The MAGIC program [54] has been used for the layout. To utilize given design rules maximally, the technology file for MAGIC has been customized so that it has a 0.1 μm minimum grid (or λ).

During the layout phase, emphasis was put on minimizing parasitic capacitances, power line inductances, and latch-up susceptibility. Parasitic capacitances were minimized by cutting drain junction areas and interconnection distances, and using minimum line widths. Also, isolating the signal lines by putting power rails in between, the coupling capacitances were reduced. Substrate and well guardrings have been extensively used to isolate individual cell from its neighbors and to prevent latch-up.

Wide second metal lanes were used for the power supply distribution, with frequent shunt contacts to the first metal. Because of this, horizontal pitch of the switch and memory cell is larger than needed for the circuit layout. This inevitably increases parasitic shunt capacitances on the rows and columns of the matrix.

Global power supply routing uses wide second metal layer. Therefore, input/output signals that are on the first metal are sandwiched between the two power supply plates: the substrate and the second metal. Even though coupling through fringe fields among signal lines is reduced, shunt capacitances increase on these lines.

5.4.2 Floor Plan

The floor plan of the chip is shown in Figure 5.18. Notice the similarity between this plan and the architecture shown in Figure 5.1. The high-speed inputs are located at the left side of the chip and the outputs are placed at the bottom and right sides. Vertical or horizontal pitches of core circuits such as the sense amplifier, the current-to-voltage converter of the input buffer, the current steering circuit of the output buffer, source address decoder, destination address latch, and timing logic are matched to the pitches of the switch/memory cell. This facilitates abutting these circuits for a compact layout. Because of the small core area, the chip size is determined by the pad pitch and the number of pads. For this prototype, the number of total pads is 48 and the chip size is 2.8 mm × 3.0 mm. The vacant space between the core and the periphery is used for power supply

distribution. Power supply lanes come from the top (V_{DD}) and bottom (V_{SS}), and form interdigitated fingers.

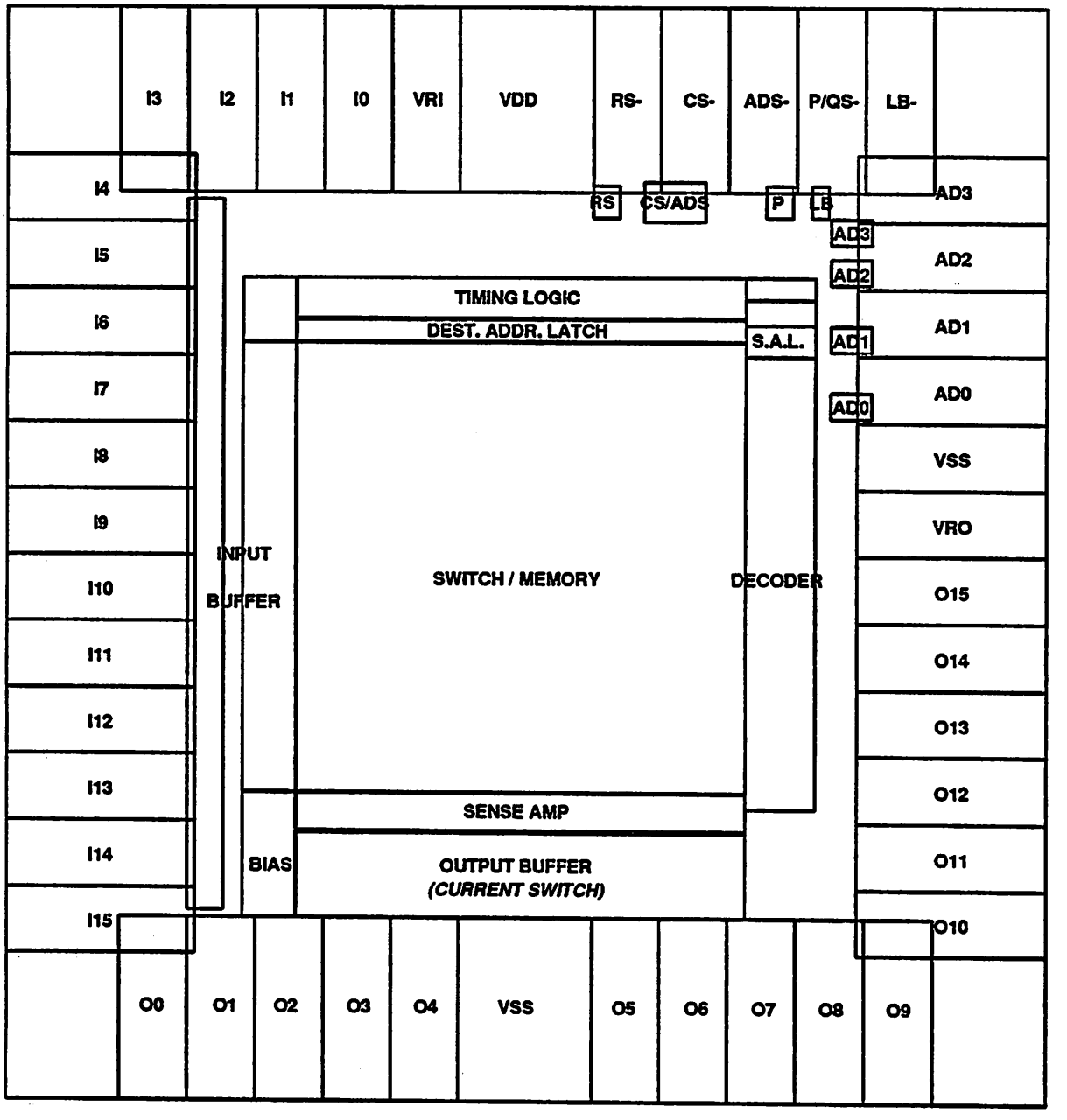


Figure 5.18 Floor plan of the chip.

The output buffer is divided into the current steering stage and the current-to-voltage converter, to minimize the effect of long routing. By placing the former right next to the switch

matrix and the latter at the pad, the steered current is delivered through the long wire. Although the shunt capacitance is large, because the voltage swing is reduced on this wire, its effect on the speed is minimized. This strategy has also been applied to the input buffer. That is, break point of the input buffer is at the output of the current steering stage. But the whole input buffer circuits are placed as close to the matrix as possible to minimize the parasitics that this buffer has to drive.

Out of the 48 pads, 5 pads are assigned to the power supplies: 3 for V_{SS} and 2 for V_{DD}. In order to reduce the power line inductances, more pads are needed for the supply connection. However, if the number of power pads is increased, the chip gets bigger and routing distances for the input/output signals become longer. This would increase the routing related capacitances and inductances, and degrades the switch bandwidth. If the second metal is used for signal routing, increase of the routing capacitances may be restricted. Because the transmission characteristics of the switch rely mostly on the power supply inductances, increasing the number of power pads would be beneficial to a certain extent.

5.4.3 Layout Verification

The layout was checked by extracting the entire circuit from the layout and doing both logic and circuit simulations on that circuit.

ESIM [54] was used to verify the logic of the control block. However, ESIM cannot handle the high-speed paths that consist of semi-analog circuits nor the pulse generation circuits that utilize gate delays.

The functioning of these circuits were checked through the circuit simulation program RELAX [55]. Actually, the voltage levels and timing for the whole chip with about 6,000 transistors have been verified from RELAX simulations. But the accuracy of these simulations is somewhat poor in this short-channel circuitry, because RELAX uses only the SPICE level 1

parameters. Also, the RELAX simulation is not reliable in that it sometimes looses fine details of transients.

More exact simulation was done for the extracted high-speed path circuits using SPICE2 and the BSIM parameters. The result is comparable to the initial SPICE2 simulation. But the bandwidth of the switch is generally lower in this case, due to over estimation of capacitances during the circuit extraction.

5.4.4 Scaled Versions

The final layout of the chip has also been scaled up by factors of 2 and 1.4, to get versions with the minimum gate lengths 2.8 μm and 2 μm and to fabricate with 3 μm and 2 μm CMOS technologies, respectively. It was done easily by using different scale factors in the CIF file generation section of the MAGIC technology file [54].

The scale-up is purely optical in a sense that the whole layout is blown up by the same factor. Hence, the pad size, pad pitch, and chip size are unnecessarily big in the scaled-up versions. In short, the scaled versions are not tuned for optimum performance. Other geometries, however, do not deviate much from the design rules of each technology, because the technologies are tailored to the scalable design rules. The chip size for the 2.8 μm version is 5.5 mm × 6.0 mm, and for the 2 μm version, it is 3.9 mm × 4.1 mm.

5.4.5 Fabrication

The 3 versions of the switch chip have been fabricated and packaged through MOSIS. The CIF files were sent to MOSIS that provides mask sets to the actual chip manufacturers. The 2 μm version has also been fabricated in the Microfabrication Laboratory of our department.

A photomicrograph of the chip is shown in Figure 5.19 which is identical for all of the 3 versions.

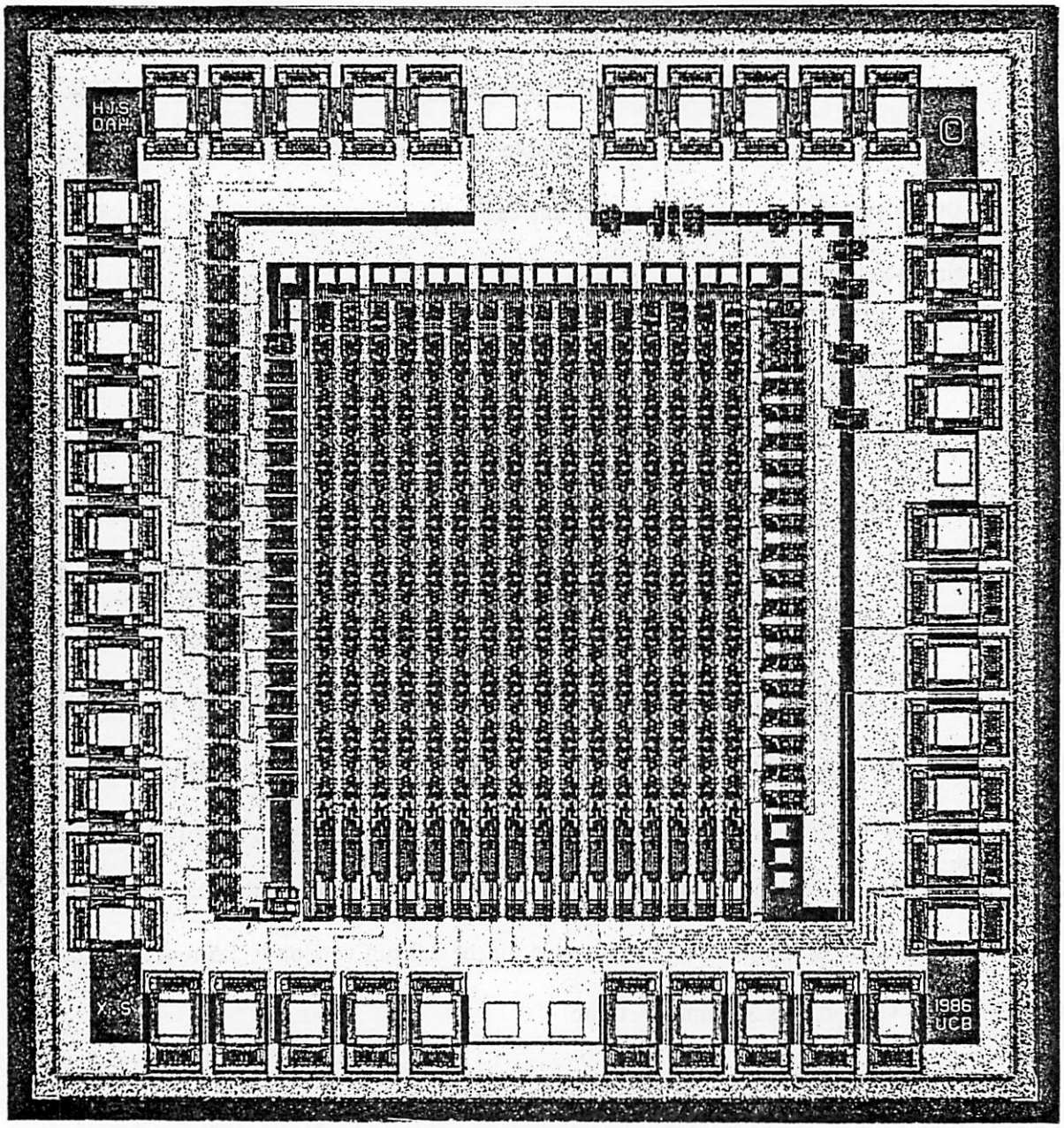


Figure 5.19 Photomicrograph of the switch chip.

48-pin LCCs were used for packaging to reduce the package lead inductance. Because typical inductance of the center lead is about 5 nH [42], the power supply inductance of this packaged

chip is approximately 2 nH. To stabilize the body bias and reduce latch-up susceptibility, the substrate is connected to V_{DD} from the back through a substrate bonding. The 2.8 μm version chips were also packaged in ordinary 64-pin DIPs.

5.5 MEASUREMENT AND DISCUSSION

The nominal characteristics of the switch chips have been measured at room temperature with a 5 V power supply. The characteristics under different operating voltages have also been observed by varying the power supply voltage from 4.5 to 5.5 V. The measurement has focused on the key characteristics of the switch such as the timing jitter, output noise, and crosstalk. Initially, the functioning of each switch was checked individually by holding the chip in a LCC socket. Then, a couple of chips have been cascaded directly to see the interface characteristics when a switching network is expanded with the tree hierarchy depicted in Figure 2.11.

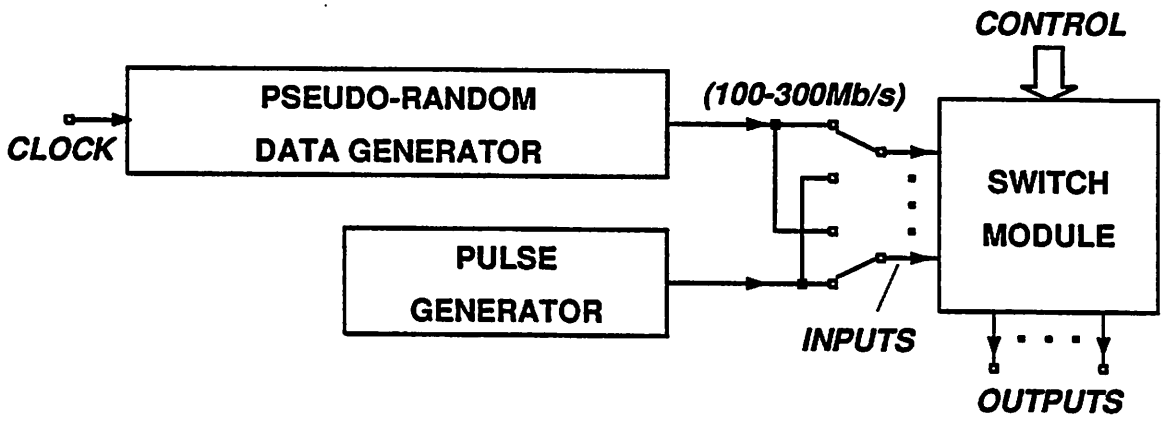


Figure 5.20 Measurement setup.

5.5.1 Measurement Setup

The measurement setup is shown in Figure 5.20. For the high-speed data inputs, a 2^{15}-1 sequence, pseudo-random NRZ data generator and a supplementary pulse generator have been

used. Oscilloscope was used to observe the waveforms or eye patterns at the input and outputs and measure the timing jitter, although the jitter would be more accurately measured if the complicated setup shown in Figure 5.21 [56] is used. The probe capacitance was 10 pF nominal.

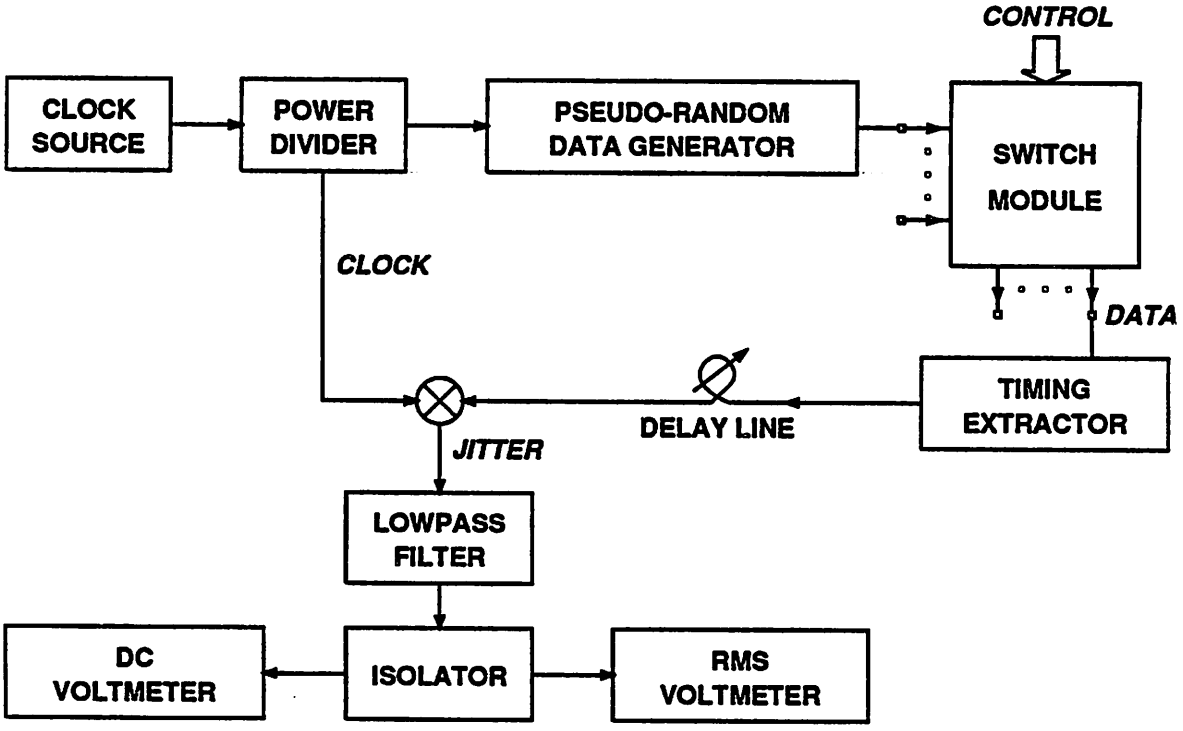


Figure 5.21 An accurate jitter measurement setup.

The pseudo-random data generator was implemented using discrete 100k series ECL components as shown in Figure 5.22. This circuit generates NRZ data up to a 300 Mb/s rate. Because the switch input is not directly ECL compatible, a level shifting circuit has been inserted at the output of the generator.

The switch module represents either a single switch in the socket or cascaded switches soldered on a custom printed circuit board. Figure 5.23 shows the connection between the switches and the top side of the board where the switches are mounted. On this board, the control port of every switch is tied to a common control bus. Therefore, when the switch is set up such that I4 is connected to O9, the input data *IN* 1 propagate through maximum 5 switches.

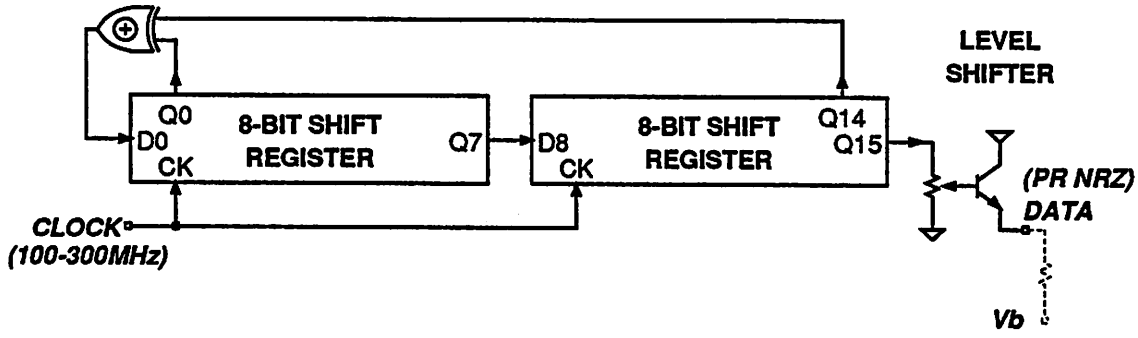


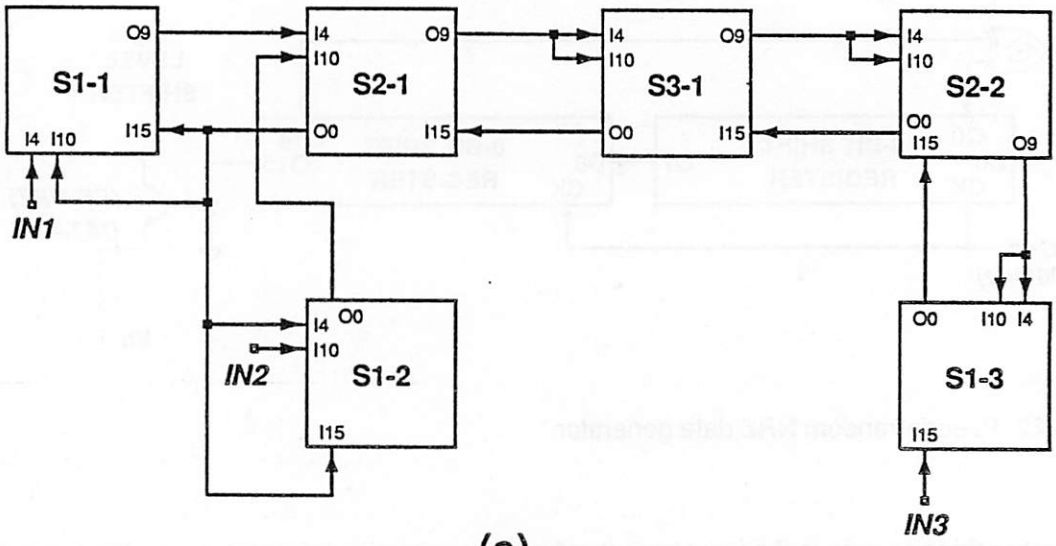
Figure 5.22 Pseudo-random NRZ data generator.

Another thing to note is the input/output reference connection between chips. For example, the switch S2-1 receives data from S1-1, S1-2, and S3-1, and the reference input node (VRI) of S2-1 must be connected to a voltage level that best approximates the center levels of the three outputs. As discussed in Appendix 2, by connecting the reference outputs (VRO) from different chips together, a reference that is close to the nominal logic threshold voltage can be obtained. The reference connection is based on this concept, and there are two reference busses for this switch module: one for VRO's of S1-1, S1-2, S3-1, and S1-3 and VRI's of S2-1 and S2-2, and the other for VRO's of S2-1 and S2-2 and VRI's of the remaining chips.

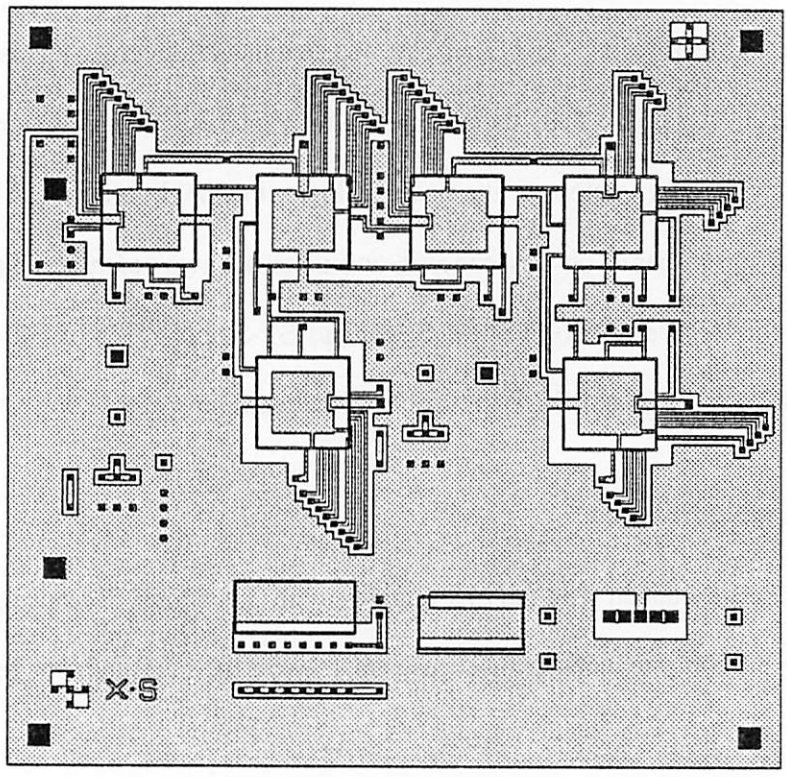
5.5.2 Measurement Results

Proper operation has been obtained with the first samples of all 3 versions. Because the 3 version chips have similar operation, this subsection focuses mainly on the measurement results for the 1.4 μm chip.

The 1.4 μm chips operate up to 300 Mb/s from a 5 V supply at room temperature. The DC output low and high levels are about 2.2 and 3.4 V, respectively. The power dissipation is about 940 mW, insensitive to whether the chip is fully broadcasting 300 Mb/s data or is in the standby state.

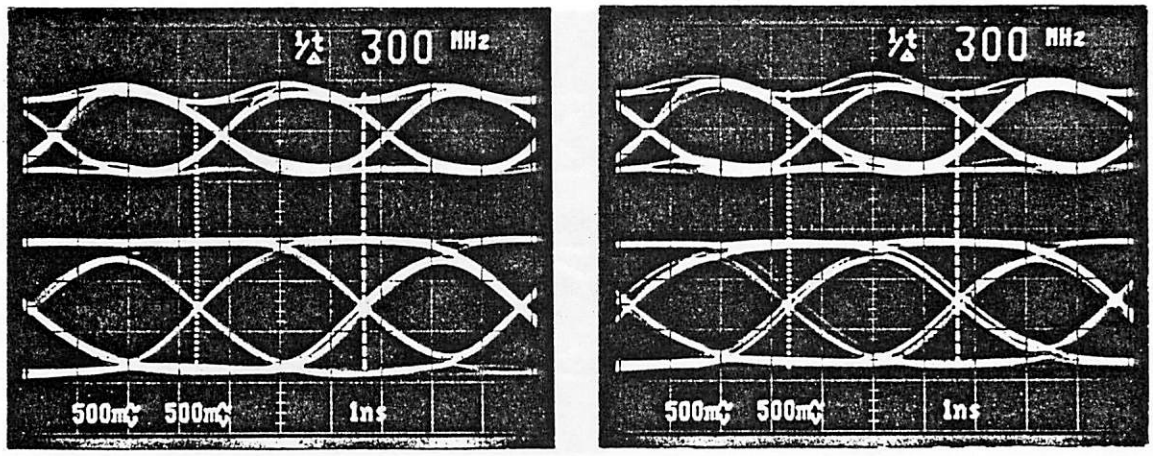


(a)



(b)

Figure 5.23 (a) Connection diagram for cascaded switches and (b) top side of the test board used.



(a) (b)

Figure 5.24 Eye patterns for the 1.4 μm chip (a) in the loop back mode and (b) in the full broadcasting mode; top traces show inputs and bottom traces show outputs.

Figure 5.24 (a) shows eye patterns of the 300 Mb/s input data (top) and the output (bottom) in the loop back mode. The output is loaded only with the probe capacitance. Note that the eye of the input data is already degraded at this data rate. In this best case, the eye is open over 80% and maximum swing is about 1.3 V. The delay is about 5.9 ns and the timing jitter added from the chip is negligible. The data rate for 80% eye-opening is 290 Mb/s. If the switch is setup for the full broadcasting mode, one input is switched to all of the 16 outputs, and on-chip power supply fluctuation and intersymbol interference become maximum. Figure 5.24 (b) shows eyes in this worst mode. Although the timing jitter is increased to about 0.3 ns and eye opening is reduced to 70%, the eye is sufficiently clear.

When the input is fully broadcast and observation is made for the output that drives another chip (fan-out = 1), the 80% eye-opening data rate decreases to 250 Mb/s, the delay increases to 6.3 ns, and the jitter is less than 0.2 ns as shown in Figure 5.25 (a). For the same condition, the data rates of 2 μm and 2.8 μm versions are 216 Mb/s and 106 Mb/s, the delays are 7.9 ns and 16.5

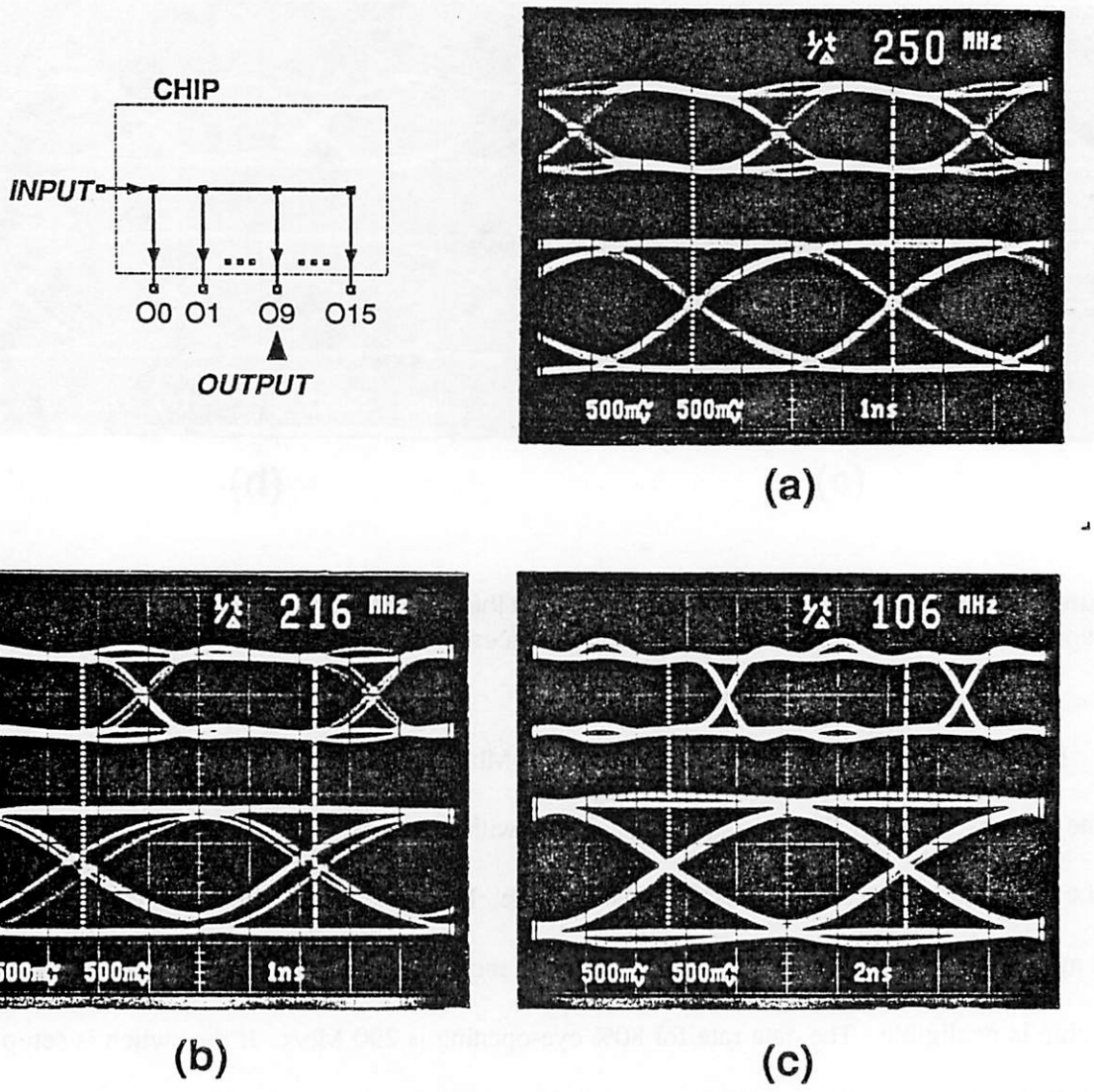


Figure 5.25 Eyes 80% open in full broadcasting mode, observed from the output with fan-out of 1: (a) 1.4 μm version, (b) 2 μm, and (c) 2.8 μm: (top) input and (bottom) output.

ns, and the timing jitters are 0.35 ns and 0.65 ns, respectively (see Figure 5.25 (b) and (c)).

One thing to note from (b) is that the output eye has two distinct lines. This suggests that the timing jitter is a manifestation of intersymbol interference which is believed to be caused by the limited bandwidth and power supply fluctuation of the switch.

Propagation delay variation has been observed between the 16 outputs of the switch in the full broadcasting mode. For the outputs with no fan-out, average delay is about 5.9 ns and the maximum deviation is about 0.2 ns (or 3%). The variation pattern is closely matched with the wire length and associated shunt capacitance between the current steering stage and transresistance amplifier stage of each output buffer. (Remember that the output buffer is divided into the two stages, and one stage is placed at the pad and the other, next to the matrix, as discussed in Section 5.4.2. Also, see the chip photograph in Figure 5.19.)

Output DC levels also vary depending on the physical location of each output pad on the chip. Low levels tend to increase and high levels decrease as the distance between the pad and the current steering stage increases. Although the level variation is minor (~4% for low level and ~2% for high level), the DC swing variation is as large as 12%. This systematic variation may come from the resistance of the wire carrying the steered output current.

5.5.3 Crosstalk Measurement

The worst crosstalk has been measured on one idle output (O9) while broadcasting the 250 Mb/s data to the remaining 15 outputs. The bottom trace of Figure 5.26 (a) shows the crosstalk noise that is superimposed on the low output level. The noise is about 140 mV p-p.

For comparison, the same output in the full broadcasting mode is shown at the top. Notice that the two eye patterns are almost identical except inversion. This inverted relationship suggests that the crosstalk may be caused through the power supply inductance coupling (on-chip supply fluctuation) or through capacitive coupling from the active outputs to the node Oy- or OCy- (see Figure 5.11) of the idle output buffer.

But close examination reveals that the latter mechanism is less probable. First of all, coupling capacitance from an output pad to the internal nodes is practically zero. Secondly, the phase difference (or delay) between the two traces is virtually negligible. If the crosstalk is

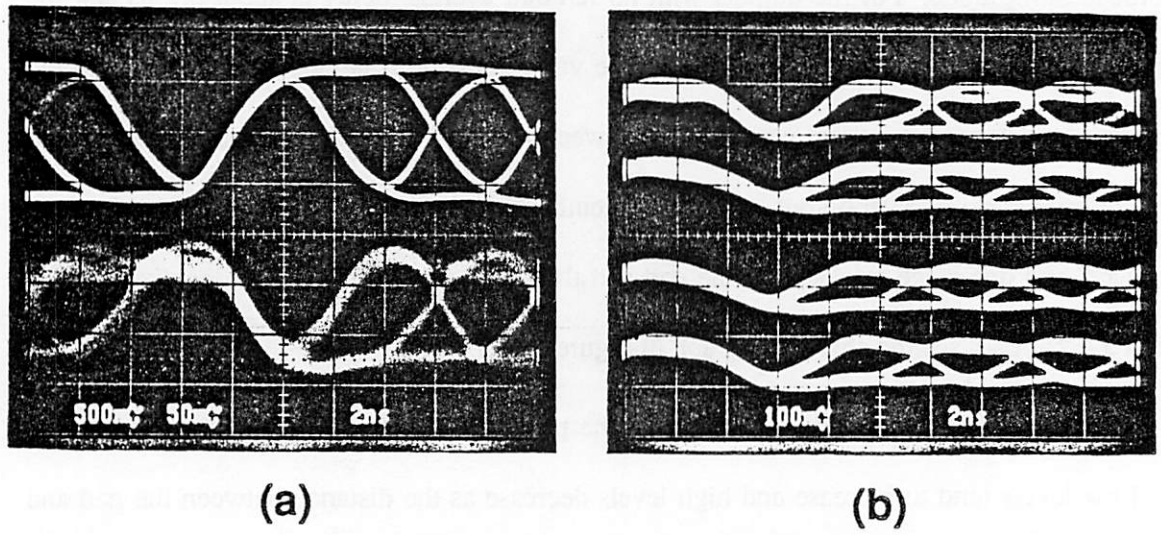


Figure 5.26 (a) 250 Mb/s data broadcast (top) and worst crosstalk observed from one idle output (bottom), and (b) crosstalk on 4 idle outputs.

caused by the latter mechanism, at least one inverter delay would be evident. Finally, the crosstalks observed from 4 idle outputs when broadcasting the same data to the rest 12 outputs are proportionately smaller (~100 mV), as shown in Figure 5.26 (b).

Therefore, the crosstalk must originate from the power supply fluctuation. In fact in the output buffer, if the output goes low, V_{SS} and V_{DD} move up, and vise versa.

For the 2 μm chip, eye pattern of 200 Mb/s data on one output affected by the maximum crosstalk that happens when broadcasting another 100 MHz asynchronous clock to the other 15 outputs is shown in Figure 5.27 at the top. Compared to the eye in the full broadcasting mode shown at the bottom, the timing jitter and eye closure are increased a little. As pointed out before, the bottom trace shows the data affected by the maximum intersymbol interference.

For the 3 μm version, the chips in LCCs show less interchannel crosstalk than the chips in 64-pin DIPs because of smaller inductances. In fact, the chips in DIPs have other forms of inter-channel crosstalk through both capacitance and mutual inductance couplings between the

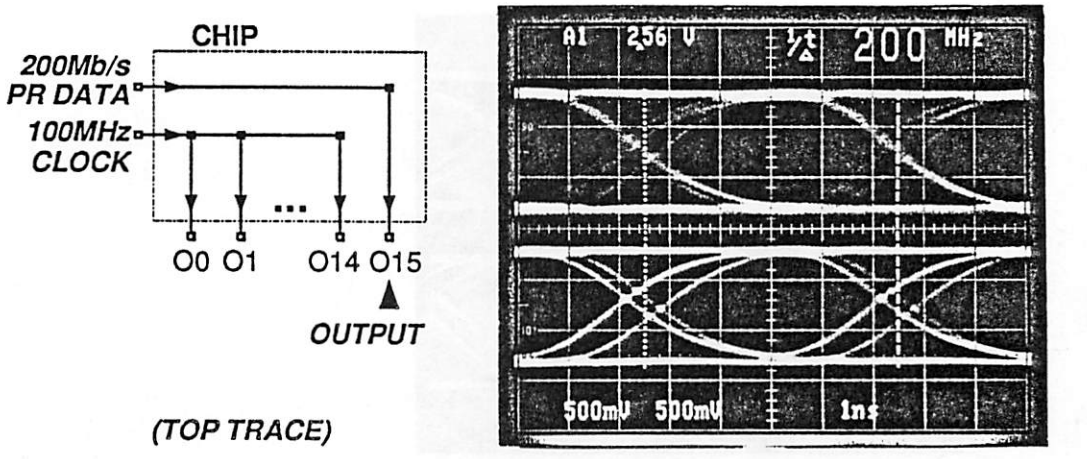


Figure 5.27 200 Mb/s data affected by maximum crosstalk (top) and maximum intersymbol interference (bottom).

package leads.

5.5.4 Expansion Experiment

Four 1.4 μm chips have been soldered on the test board shown in Figure 5.23 (S1-1, S1-2, S2-1, and S3-1), to see expansion related interface characteristics. The switches are set up for full broadcasting, and 250 Mb/s data are sent from the input of S1-1.

In this worst mode, the data propagate through the 4 chips. Figure 5.28 shows the output O9 of S2-1 (fan-out = 1) at the top and O9 of S3-1 (fan-out = 0) at the bottom. The output of S1-2 is not shown but similar to the bottom trace. Compared to the trace shown in Figure 5.25 (a) which is in fact O9 of S1-1, the eyes become worse as the data pass through more chips.

This increase of timing jitter mainly comes from the offset between the input reference and center level of the input to each chip. As described in Section 5.5.1, the reference input terminal VRI of each chip is connected to the reference outputs of the adjacent chips that are shorted together. Again, because of the electrical parameter variation from chip to chip, the reference

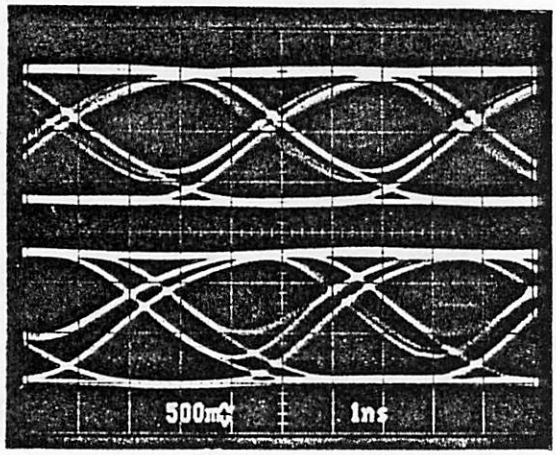


Figure 5.28 Outputs of S2-1 (top) and S3-1 (bottom) show successively closing eyes when data propagate 2 cascaded chips and 3 chips.

outputs provided as the logic threshold voltages are different between chips. (For the 1.4 μm version, the reference output voltage is typically 2.78 V.) If several reference outputs are shunted, the resulting level becomes close to the average of them and would not represent the center level of any output swing (Appendix 2). Another thing to note is that the outputs from one chip may also have small center level variation. Therefore, for example, VRI of S2-1 may not match the center level of the input from S1-1 (O9).

Figure 5.29 illustrates this mismatch between the input center level and the reference level, and its effect on the data propagation. If the pulse transition edges are infinitely abrupt, the level mismatch does not cause any major problem. However, outputs of the switch have pulse transition periods (i.e., rise and fall times) determined by the bandwidth of the switch. As a result, the decision points and effective pulse widths vary for different inputs. Especially, when the data rate is near the maximum capacity of the switch, the pulse transition period is approximately equal to the bit period, and the decision point variation severely affects the propagation of data through the switch. In other words, effective duty cycle of the input would deviate from 50%

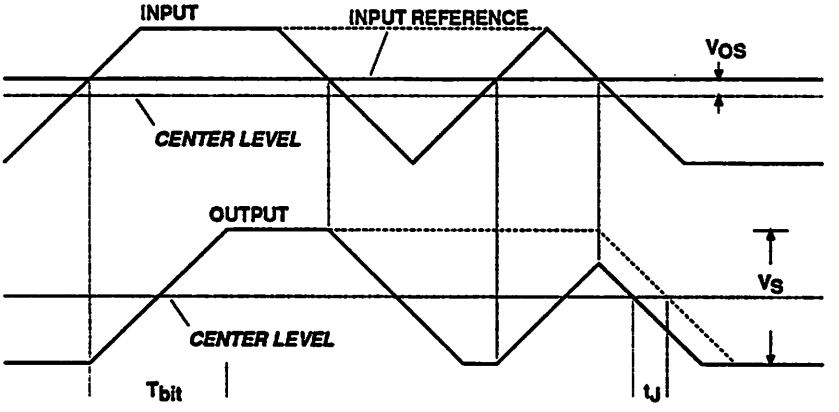


Figure 5.29 Offset between the input reference and the center level of the input causes timing jitter.

resulting in one bit period shorter than the switch capacity. Therefore, the data may not propagate properly or may have an excessive timing jitter. In this worst case when one bit period is equal to the transition period, assuming that the input transitions are linear, the maximum jitter added from a level offset V_{OS} becomes

$$t_J = \frac{2V_{OS}}{V_S} T_{bit}. \tag{5.4}$$

where V_S is the input swing and T_{bit} is the bit period. For example, a 50 mV offset for a 1 V swing causes 10% timing jitter. Of course, if the pulse transition edges are abrupt compared to the bit period, the jitter is minor.

For the 2.8 μm version, the outputs after many chips have been examined using a spectrum analyzer. The frequency spectra become more rough and jagged, but jitter addition could not be estimated from the data.

5.5.5 Power Supply Voltage Sensitivity

Sensitivity of the chip characteristics on power supply voltage variation is measured for the 3 versions. For a ± 10% supply variation, i.e. from 4.5 V to 5.5 V, Table 5.1 summarizes the charac-

teristics change of the 1.4 μm chip. Similar behavior has been observed for the other 2 scaled versions.

SUPPLY VOLTAGE [V]	4.5	5.0	5.5	(20%)
SUPPLY CURRENT [mA]	153	188	228	(40%)
OUTPUT SWING [V]	1.11	1.23	1.33	(18%)
PROPAGATION DELAY [ns]	6.78	6.34	6.12	(10%)
DATA RATE [Mb/s]	244	250	254	(4%)
CROSSTALK [mV]	137	139	145	(6%)

Table 5.1 Characteristics of the 1.4 μm chip for different power supply voltages.

The power supply current that is mainly static varies about 40%. This is expected because major portion of the current flows in the transresistance amplifiers that are biased at the logic threshold voltage. In this case, the current is roughly proportional to $(\frac{V_{DD}}{2} - V_{Tn})^2$.

One thing to note is the output swing variation (~18%) or the sensitivity of the swing that is about 0.9. This sensitivity lower than 1 verifies the effectiveness of the bias generator and feedback resistor combination discussed in Sections 5.2.4 and 4.4.7.

Other things to note include the delay and the data rate for 80% eye opening in the full broadcasting mode. Although the delay decreases and the data rate increases as the supply voltage is increased, they show unexpectedly low variations: 10% and 4%, respectively. This suggests that a large part of the propagation delay may not be related to charging capacitances. If the high-speed path circuits are simple digital inverters, the delay and data rate variation would be

close to 20%.

The crosstalk that would depend on the output swing and data rate, varies only 6%. This is far less than expected (~22% for given data) and suggests that the transient current sensitivity is quite low.

5.5.6 Performance under Scaling

The chip characteristics of the 3 versions are summarized in Table 5.2. The 3 versions can roughly be considered as designs following the constant voltage scaling law [44] with a scaling factor, S, of 1.4. Under the constant voltage scaling, the horizontal and vertical dimensions are scaled but the power supply voltage is kept constant.

Versions	*2.8 micron Chip*	*2 micron Chip*	*1.4 micron Chip*
Chip Size	5.5 mm x 5.9 mm	3.9 mm x 4.1 mm	2.8 mm x 3.0 mm
Power Dissipation	0.8 W	1.3 W	0.9 W
Max. Data Rate	130 Mb/s NRZ	260 Mb/s NRZ	300 Mb/s NRZ
Propagation Delay	15.5 ns	7.6 ns	5.9 ns
Max. Timing Jitter	9% @105 Mb/s	8.5% @215 Mb/s	4% @250 Mb/s
Max. Crosstalk	110 mV p-p	100 mV p-p	140 mV p-p

Table 5.2 Characteristics of the 2.8, 2, and 1.4 μm chips for a 5 V supply at room temperature.

Thus the chip size is expected to decrease by S^2 going from the 2.8 μm version to the 2 μm version and from the 2 μm to the 1.4 μm, which is evident from the table.

The power dissipation that is proportional to the static supply current is expected to increase by S because the gate oxide thickness is reduced by S and k' increases by S. However, the dissipation actually increases by 1.6 for the 2.8 μm - 2 μm scaling and 0.7 (decrease!) for the 2 μm -

1.4 μm scaling. This means that the 3 technologies are not scaled properly in terms of the current. In fact, k' of the minimum gate length n-MOSFET in each technology does not follow the scaling law. As shown in Table 5.3, it increases by factors of 1.2 and 1.2. The deviation of scale factors is worse for the gate oxide thickness: 1.7 and 1.1. Although the trend of gate oxide thickness reduction is close to the trend of power dissipation, the discrepancy is still large. The reasons may include excessive lateral diffusion that reduces effective channel length for the 2 μm version, the short channel effects that reduce saturation current for the 1.4 μm version, and inaccurate modeling that affects parameters such as k'.

Versions	*2.8 micron*	*2 micron*		*1.4 micron*
tox (NMOS)	1	$\frac{1}{1.7}$	1	$\frac{1}{1.1}$
k' (NMOS)	1	1.2	1	1.2
Power Dissipation	1	1.6	1	0.7
Delay	1	$\frac{1}{2.0}$	1	$\frac{1}{1.3}$
Data Rate	1	2.1	1	1.2
Crosstalk	1	0.9	1	1.4

Table 5.3 Actual scale factors for the parameters and chip characteristics.

The delay is expected to decrease by S^2 for a perfect scaling. But as the oxide thickness and current scale differently, the delay may scale according to

$$S_D = \frac{S^2 S_I}{S_{tox}} \tag{5.5}$$

where S_{tox} is the oxide thickness scale factor and S_I is the current scale factor. If the actual values are used, the delay scale factors in the table are verified.

Also, because the data rate is inversely proportional to the delay, the same scale factors are expected for the data rate. This expectation is proved in Table 5.3.

The crosstalk most of which comes from the transient current in the output buffer may depend on the power dissipation and the data rate. That is, the crosstalk increases by S^3 or, in practice, by $S_I S_D$. The data shown in the table, however, deviate from the prediction too much. This would be due to missing of the substrate bonding for the 2.8 μm chip and inaccuracy of the measurement.

5.5.7 Conclusion

The measurement results suggest that the circuit techniques used for the prototype switch reduce the effects of chip parasitics, especially power supply line inductances, and allow the propagation of high-speed (up to 300 Mb/s) data. Timing jitter added from the switch is less than 10% and on-chip power supply noise appeared as the crosstalk is around 100 mV (~10% of the output swing) at the 80% eye opening data rate.

The scaling experiment demonstrates that the design is robust against blind scaling. Although the circuit has not been optimized for the 3 μm or 2 μm technology, proper operation has been obtained from the scaled-up versions and performance of those is quite close to what is expected.

The circuit techniques can also be applied to other high-speed circuits such as sense amplifiers in fast static or dynamic memories. In addition, they may be applicable for reducing voltage swings on the highly capacitive, high-speed data busses in processors. Without increasing device sizes and without introducing power supply noise, desired speeds will be achieved by using the techniques.

5.6 FUTURE IMPROVEMENTS

The prototype chip described in this chapter consumes unnecessarily large static power in the high-speed paths. Also it has minor interface problems when constructing a large switching network. This section suggests possible improvements to these problems.

5.6.1 Reduction of Static Power Consumption

As has been pointed out, most of the static power is consumed in the transresistance amplifier circuits. Among the three transresistance amplifiers along the high-speed path, the one in the output buffer dissipates about 2/3 of power. Remember that, during the standby state when none of the 16 switch cells that are connected to one output channel is ON, the output is pulled down to a low level (see Figures 5.5, 5.6, and 5.11). This standby low level is basically identical to the lower limit of the swing for an active output. Because of this, guard time required between each data transmission is minimized. But in applications where enough guard time is provided, the standby low level may be pulled down to V_{SS} so that static power consumption is reduced.

One way to pull down the output low level to V_{SS} is to break the feedback path in the output transresistance amplifier during the standby state. This can be achieved by controlling the transmission gate feedback resistor. Since the standby pull-down signal SBy is already available, it can be used for the control. Figure 5.30 illustrates this concept; SBy is connected to the gates of p-MOSFETs MPA2 and MPO5 and its complement is connected to MNA2 and MNO5. Therefore when SBy is high, the transmission gates are OFF and Oy and OUTy are discharged to V_{SS}. This modified circuit eliminates static power dissipation in the sense amplifier and output buffer during the standby state.

5.6.2 Output Level Control

As discussed in the previous section, constructing a large switch network by the method illus-

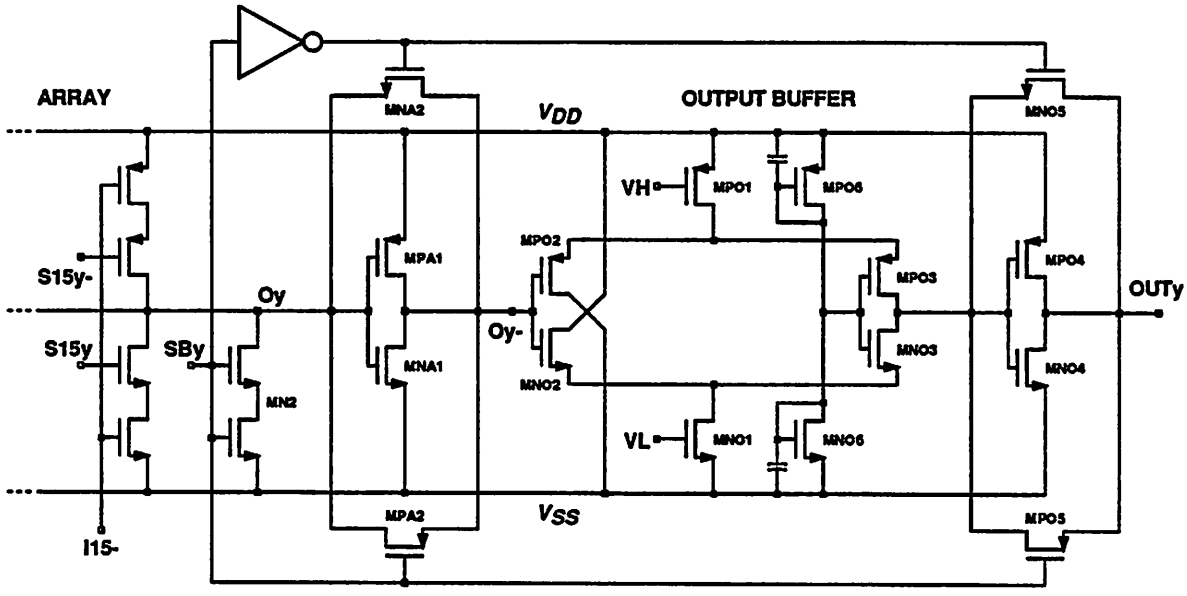


Figure 5.30 Sense amplifier and output buffer modified for minimum static power consumption.

trated in Figure 5.23 (or Figure 2.11 and 2.12 in general) degrades the effective network bandwidth, because the level mismatches between outputs from different chips change the duty cycles of pulses and reduce bit periods.

One solution to this problem is to provide one input reference for the whole switch network and control the output center levels as close to this reference as possible. The center level of the output buffer shown in Figure 5.11 can be controlled by adding a transistor as shown in Figure 5.31. In fact, the logic threshold voltage of the last inverter changes if the conductance of MNO7 is varied through Vcont. The voltage Vcont needs to be controlled so that the final logic threshold voltage is very close to the input reference.

The control voltage can be generated using a dummy high-speed path circuit as shown in Figure 5.32. The negative feedback loop forces the logic threshold voltage OUTd to be equal to the reference input VR and generates a proper Vcont. Simulation shows that the logic threshold tracks the input well between 2.2 and 3 V with a little offset ~15 mV (Figure 5.32 (b)).

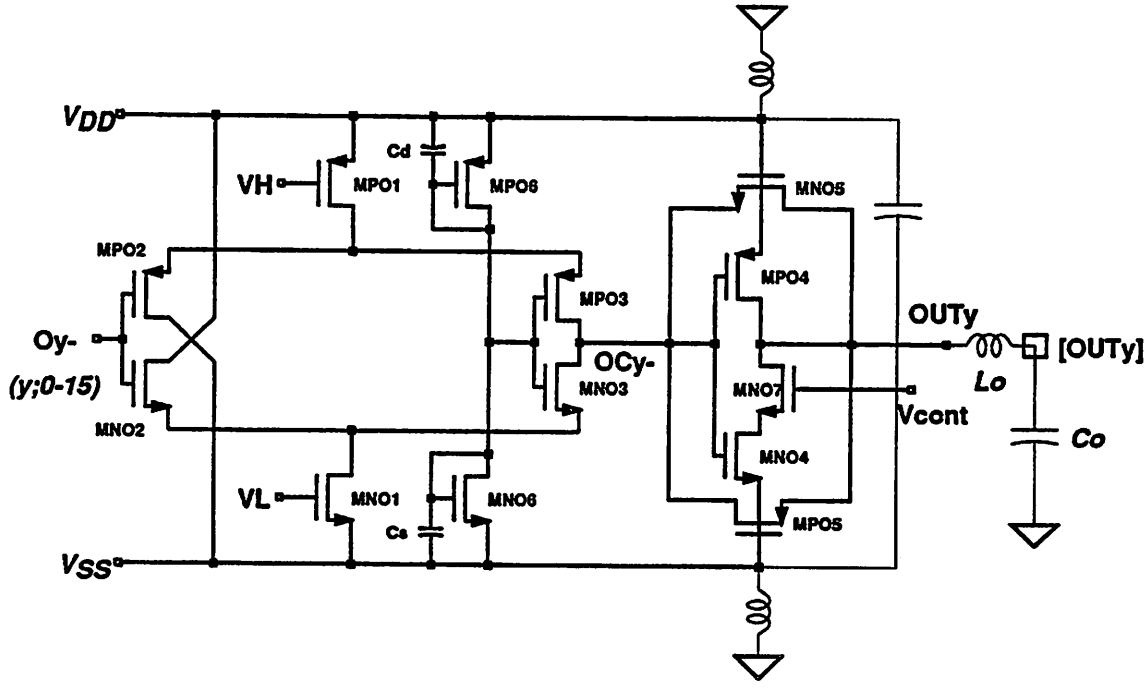


Figure 5.31 Output buffer modified for output level control.

Although the output levels in one chip follows the input reference in general, due to parameter variation at different locations, the levels of different outputs would have non-identical offsets. Another thing to note is that the parameter variation is much more severe among different chips. However, since this circuit ensures that the output levels of an individual chip track the reference with minor offsets, the chip-to-chip interface is much better than that for the prototype switch.

5.6.3 Current Interface

The interface between different chips can also be improved by designing the output buffer current driving and input buffer current sensing. Figure 5.33 shows one design modified from the prototype switch. In this circuit, the voltage swing reduction technique is applied at the highly capacitive output node that is directly driven by the switch cells. The transresistance amplifier in the input buffer senses current from the switch and limits voltage swing at the off-chip node.

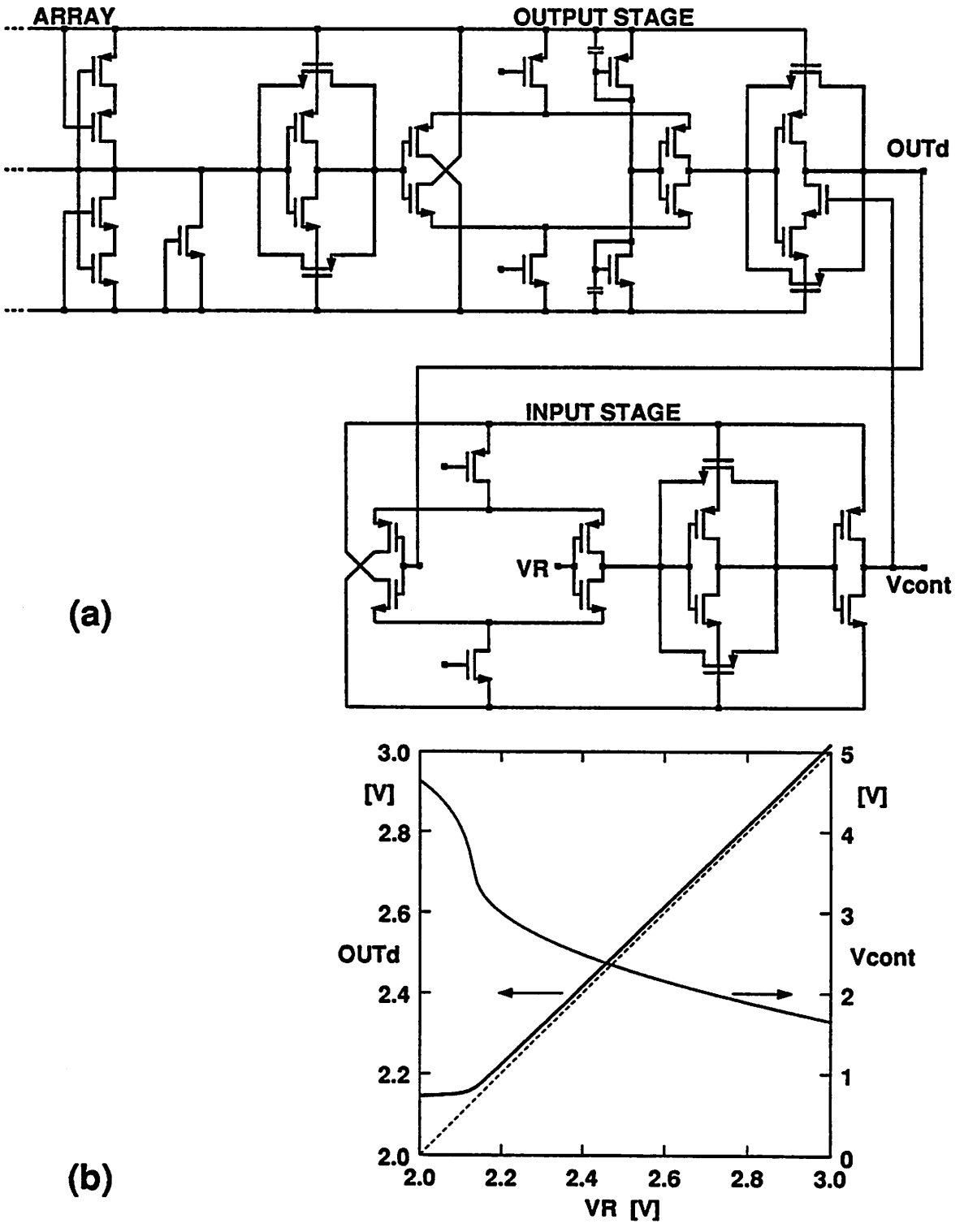
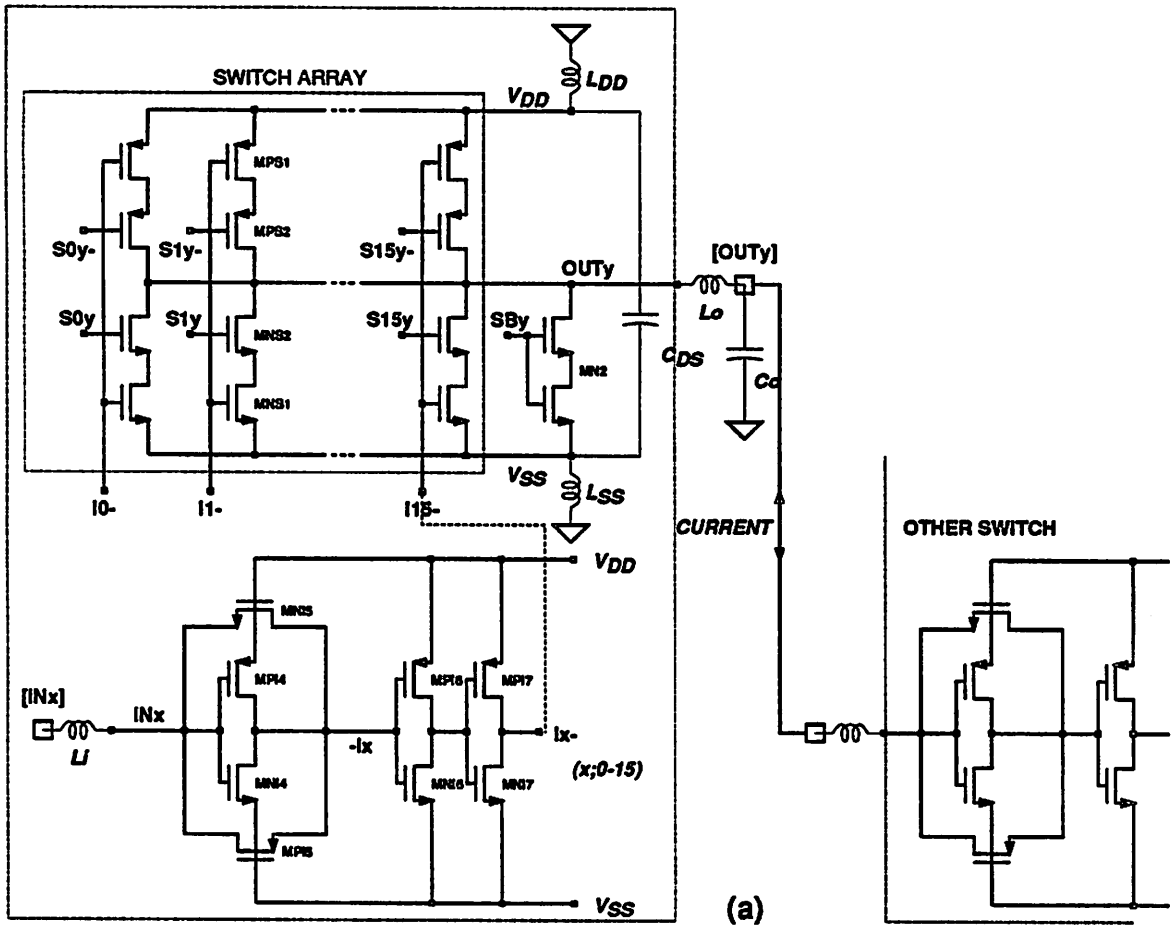


Figure 5.32 (a) Dummy circuit for generating control voltage and (b) tracking performance of the logic threshold voltage to the input reference.



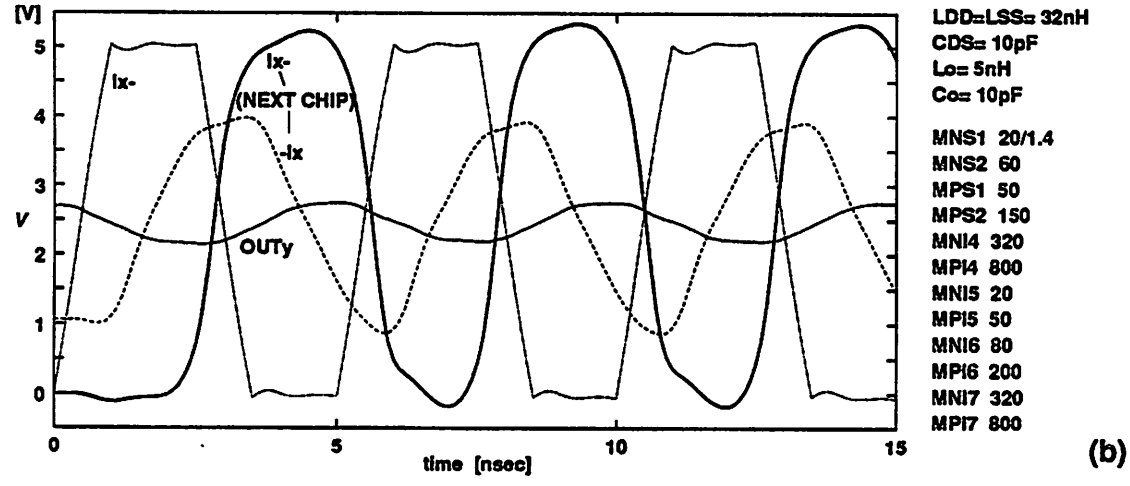


Figure 5.33 (a) Current-driving output buffer and current-sensing input buffer and (b) worst-case simulation results of the circuit.

This design is advantageous because the input/output interface is insensitive to parameter variation between the chips and the signal path from input to output is simplified. A worst case simulation shown in (b) suggests that 400 Mb/s NRZ data propagate with about 2.5 ns delay and 4% timing jitter.

One drawback of this design is that the output can only drive one input. In other words, the fan-out is limited to 1 in normal operations, which restricts the applications of the chip in general. However, because most of the expansion methods require switches with the fan-out of 1 (see Figures 2.9 - 2.12), this design would be sufficient for most applications. Another disadvantage is the large power dissipation in the input buffer. This is because current from the switch cell needs to be large to drive the off-chip capacitive load in short time, and the current in the transresistance amplifier must be larger than this current for proper operation.

CHAPTER 6

CONCLUSION

This thesis concludes that wideband circuit switching for 100 to 200 Mb/s fiber-optic networks is feasible in prevalent, low power, large-scale silicon CMOS IC technologies, even though not the fastest. This suggests that the future fiber based LAN or broadband ISDN would not require underdeveloped and expensive optical techniques. Optical domain channel switching would be needed only at the highest data rates on the order of 1 GHz.

Electronic switching has been demonstrated by implementing a 16×16 crosspoint switch, which is one of the key elements in circuit switching networks, using 3 to 1.25 μm CMOS technologies along with appropriate circuit techniques.

Circuit techniques suggested in this work reduce voltage swing and current transients at critical nodes, enhance the switch bandwidth, and allow integration of more crosspoints for a given technology while reducing power line noise, crosstalk, and intersymbol interference. Furthermore, the switch is directly scalable and the effectiveness of the techniques is maintained under technology scaling.

These circuit techniques would be applicable to the CMOS realization of other wideband circuits such as opto-electric receivers, transmitters, and timing recovery and regeneration circuits that are needed for broadband electronic switching, and to high speed digital CMOS ICs, in general.

APPENDIX 1

OPTIMAL DESIGN OF DRIVER CHAINS

There are many nodes inside MOS VLSI digital circuits that have large capacitance, such as clock lines, data busses, control lines for parallel data paths, inputs of large decoders, and output pads. To drive such nodes without introducing excessive signal delay, a large amount of current needs to be sourced from the driving circuit. But the drive required for these heavy loads is beyond the capability of the small gates that are used for ordinary logic implementation in VLSI chips. In this case, a driver chain that is a cascade of drivers whose sizes are successively scaled up can be inserted between the minimum size gate and load. Often in CMOS digital circuits, a driver is simply realized with a static inverter. However, a driver chain needs to be optimized in terms of the number of stages and scale up factor to minimize the propagation delay. In fact, the optimal criteria for a driver chain have been derived in view of minimizing the delay [57].

The previous derivation lacks certain details that may become important in real implementations. Each driver in the chain has parasitic capacitances at its output node and, due to these parasitics, the delay through that driver is not purely proportional to the input capacitance of the next stage as has been assumed in the derivation. This can result in substantial errors in the optimum values. Another thing to note is that the delay is not the only performance figure to be optimized in VLSI circuits; power and area are also important. Optimization without considering these performance criteria will lead to circuits with wasteful die area and power consumption. Especially since optimal design guides are being set up for computer-aided design of VLSI circuits, even small errors in the guides may be detrimental.

In this appendix, the effect of parasitics is addressed and the optimum scale up factor that

minimizes the total propagation delay through a general driver chain is derived. Optimization based on area-delay, power-delay, and area-power-delay products as alternative figures of merit are also discussed.

A1.1 DRIVER CHAIN DELAY OPTIMIZATION

An N-stage driver chain driving a large capacitive load C_L or sourcing a large current I_N is shown in Figure A1.1. The first stage of the chain is driven by a current I_0 from a minimum size gate whose input capacitance is C_{G0}. The i^{th} driver sources an average output current I_i and has an input capacitance C_{Gi}. And the size of each driver in the chain is increased by a scale up factor κ from that of the previous stage. This means $I_i = \kappa I_{i-1}$ and $C_{Gi} = \kappa C_{Gi-1}$. In fact, this successive scale up with a constant factor ensures a minimum delay when interconnection capacitances between the stages are negligible [58]. Therefore, if η is defined as C_L/C_{G0},

$$\eta = \kappa^{N+1}. \tag{A1.1}$$

Since each stage has parasitic capacitances associated with diffused junction, gate overlap, and interconnection at the output node, the propagation delay through the stage is not linearly proportional to the next-stage input capacitance. Denoting the total parasitic capacitance of the i^{th} driver as C_{Pi}, the delay for driving the capacitance load is proportional to $(C_{Gi+1} + C_{Pi})/I_i$ or $(\kappa C_{Gi} + C_{Pi})/I_i$. Because the delay due to the capacitance dominates over the intrinsic device transit time in most cases, the propagation delay at the i^{th} stage is represented as

$$t_{Di} = \frac{\kappa C_{Gi} + C_{Pi}}{C_{Gi} + C_{Pi}} \tau, \tag{A1.2}$$

where τ is the time delay through a driver driving another same size driver ($\kappa = 1$). For an inverting driver, τ is the familiar propagation delay of a technology, usually measured at fan in = fan out = 1 using ring oscillators. Note that the propagation delay at each stage depends on C_{Pi}.

Therefore, if C_{Pi} is random from stage to stage, the optimization problem gets complicated.

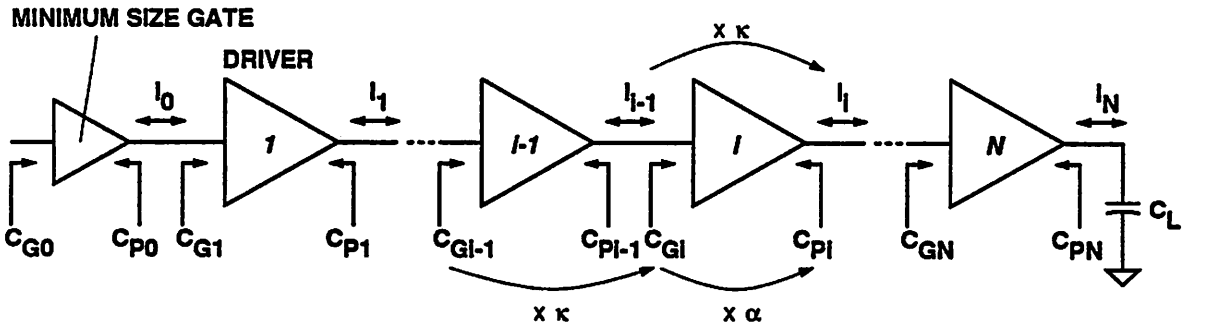


Figure A1.1 N-stage driver chain driving a highly capacitive load or sourcing a large current.

However, the gate overlap capacitance is exactly proportional to the widths of MOSFETs used in a driver. In addition, because more contact area is needed for larger drive, the diffused junction area and corresponding capacitance at the output are approximately proportional to the transistor widths. Also in practice, the two adjacent driver stages are closely located and the interconnection capacitance between them is negligible. Hence, there exists a proportionality α between C_{Pi} and C_{Gi} that is determined by the actual layout and technology and remains unchanged throughout the driver chain: $\alpha \equiv C_{Pi}/C_{Gi}$. Furthermore, since the drive capacity is controlled by the MOSFET widths, the junction and gate overlap capacitances scale by the same factor κ between stages. As a result, the delay time per stage becomes

$$t_{Di} = t_D = \frac{\kappa + \alpha}{1 + \alpha}\,\tau. \tag{A1.3}$$

Note that, for $\alpha = 0$, t_D reduces to $\kappa\tau$ or is linearly proportional to κ.

The total delay t_{TD} through the chain simply sums to $(N + 1)t_D$. Combining Eqs. (A1.1) and (A1.3),

$$t_{TD} = \frac{(\kappa + \alpha)\ln(\eta)}{(1 + \alpha)\ln(\kappa)}\,\tau. \tag{A1.4}$$

Therefore, the total delay is proportional to the technology dependent parameter $\tau/(1 + \alpha)$ and to

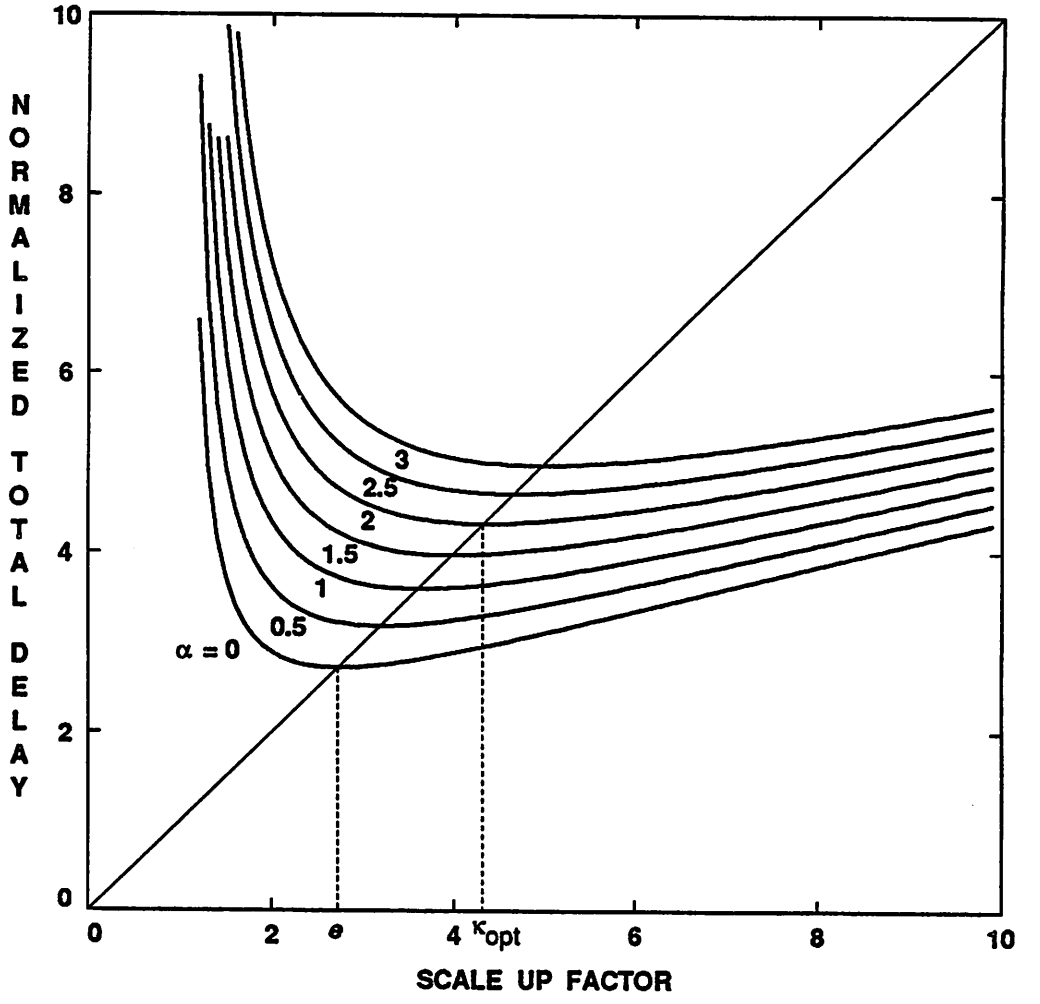


Figure A1.2 Normalized total delay of a driver chain versus scale up factor κ for different parasitic capacitance ratios α.

the logarithmic capacitance ratio $\ln(\eta)$. It is also a function of the scale up factor; the function $(\kappa + \alpha)/\ln(\kappa)$, as normalized total delay t_{TDn}, is plotted in Figure A1.2 with α as a parameter. Each curve shows a unique minimum at κ_{opt} and relative flatness to the right of the optimum point. From $\partial t_{TDn}/\partial\kappa = 0$, the minima occur at the intersections of the curves with a line function κ, i.e.,

$$\kappa_{opt} = \frac{\kappa_{opt} + \alpha}{\ln(\kappa_{opt})}. \tag{A1.5}$$

When $\alpha = 0$, κ_{opt} and the normalized minimum delay have the same value e that is the base of natural logarithm, as before [57]. For other values of α, κ_{opt} can be roughly approximated by

$e + 0.8\alpha$. Using κ_{opt}, the minimum total delay becomes

$$t_{TD\min} = \frac{\ln(\eta)}{1+\alpha}\,\kappa_{opt}\tau. \tag{A1.6}$$

Another optimization based on the current ratio I_N/I_0 is possible but more complicated.

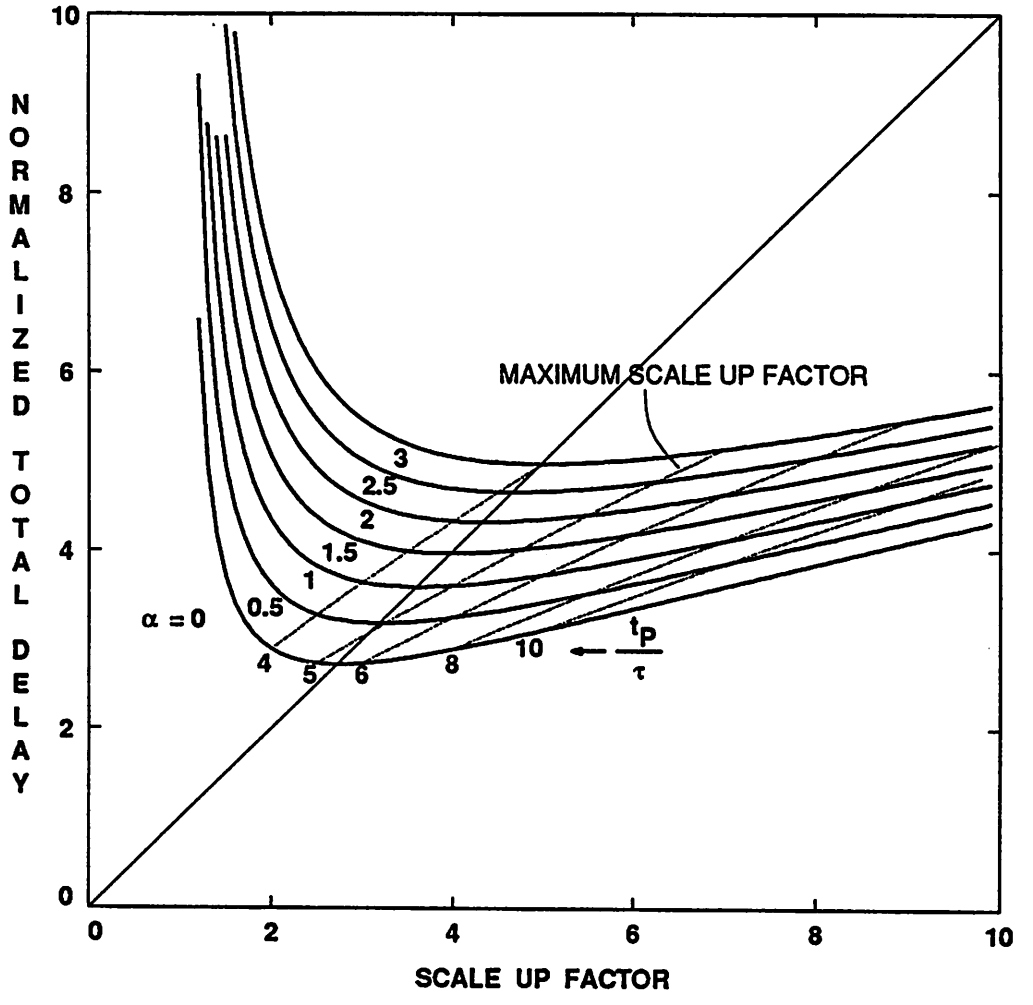


Figure A1.3 Normalized total delay versus scale up factor for different parasitic capacitance ratios constrained by Eq. (A1.7).

The optimization shown above is somewhat impractical in a sense that it ignores the following constraints; N must be an integer and t_D be less than half the minimum pulse period ('1' or '0'). To satisfy the first constraint, simply get κ_{opt} and calculate N. Then select two integers closest to N and, for each, obtain κ. By comparing the corresponding total delays, the real

optimum values of κ, N, and t_{TD} can be obtained. Generally, the smaller of the two integers would be the optimum one because the delay curves are relatively flat for a κ larger than κ_{opt}.

The second constraint comes from the fact that, if the propagation delay through a driver is only due to the output capacitance, rise or fall time is twice the delay time and the rise or fall transient has to be completed before the opposite transient starts. This condition can be restated as, using Eq. (A1.3),

$$\kappa < \frac{(1+\alpha)}{2}\frac{t_P}{\tau} - \alpha, \tag{A1.7}$$

where t_P is the minimum pulse period. Normally the minimum pulse period is half the clock period of a circuit. The resulting restriction is superimposed on the delay curves as shown in Figure A1.3 for different t_P/τ values. Note that, for large t_P/τ values, the above optimization applies well. However, for values smaller than 5, the curves begin to exclude optimum points. This situation would happen when trying to design high-speed circuits out of a limited bandwidth technology. In that case, the minimum delay is obtained for the maximum κ.

A1.2 AREA AND POWER MINIMIZATION

Area of each driver in Figure A1.1 is approximately proportional to the drive capacity, because both are linearly related to the transistor widths. Thus, if the area of the minimum size gate is A_0, that of the i^{th} driver becomes $A_0\kappa^i$. The total area used by the driver chain will be

$$A_T \equiv A_0\sum_{i=1}\kappa^i = \frac{\kappa^{N+1}-\kappa}{\kappa-1}A_0 = \frac{\eta-\kappa}{\kappa-1}A_0, \tag{A1.8}$$

which decreases steadily as the scale up factor is increased. One reasonable optimization parameter would be the delay-area product of the total chain. The normalized total delay-area product,

$$DA_{T_n} = \frac{(\kappa+\alpha)(\eta-\kappa)}{(\kappa-1)\ln(\kappa)}, \tag{A1.9}$$

is plotted in Figure A1.4 for η = 10, 100, 1000 and α = 0, 1, 2. The delay-area product is seen to be a monotonically decreasing function of the scale up factor. So, to minimize the delay-area, it is preferable to select a scale up factor as large as possible unless the delay gets out of a specified limit.

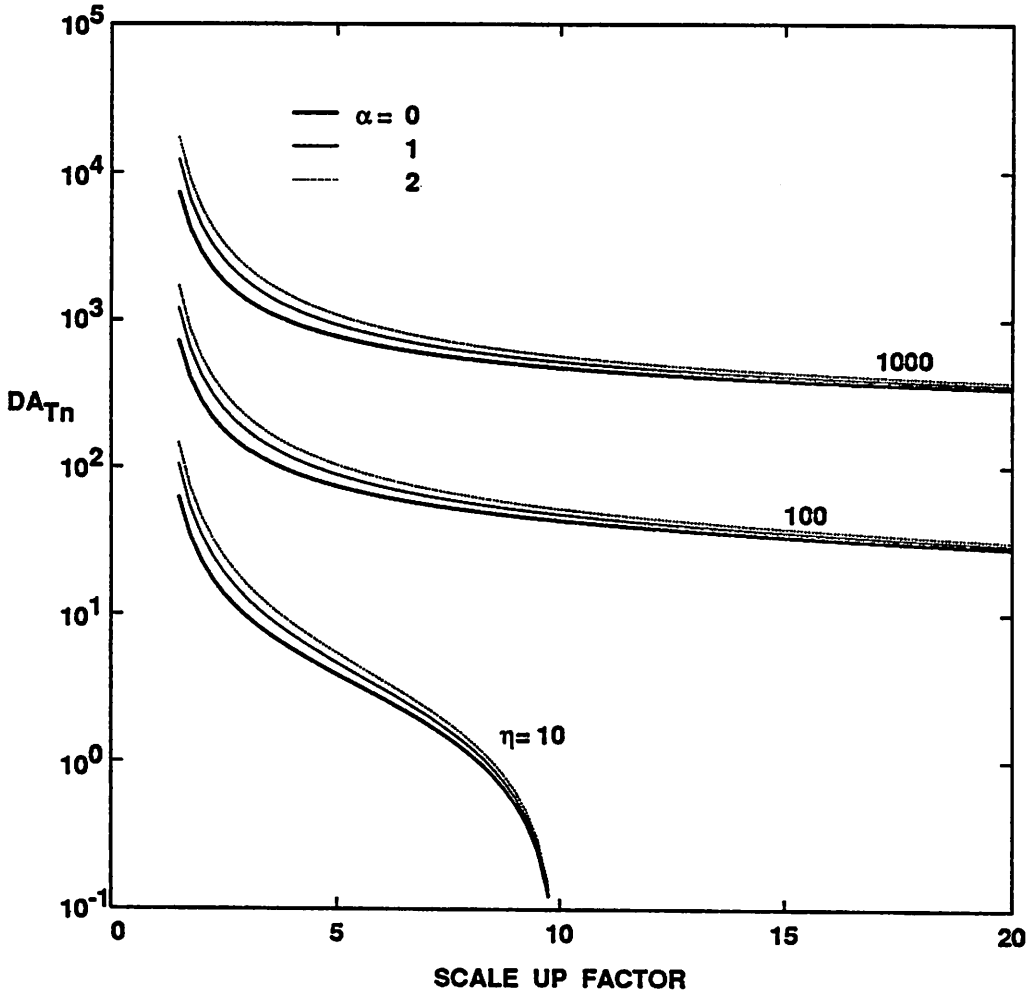


Figure A1.4 Normalized total delay-area product versus scale up factor for different capacitance ratios, η and α.

Power consumption of each driver, whether static or dynamic, is also proportional to the drive capability. This is obvious for static power and can be deduced for dynamic power. The dynamic power at the output of the i^{th} driver is expressed as $C_i V_i^2 f$, where $C_i = C_{Gi+1} + C_{Pi}$, and

V_i is the voltage swing. Assuming that $V_i = V$ and the static power is negligible in the driver chain, the total power consumption becomes

$$P_T = V^2 f \sum_{i=0}(C_{Gi+1} + C_{Pi}) = C_{G0} V^2 f \sum_{i=0}(\kappa + \alpha)\kappa^i = C_{G0} V^2 f \ (\eta - 1) \frac{\kappa + \alpha}{\kappa - 1}. \qquad \text{(A1.10)}$$

The normalized total power $P_{Tn} = (\eta - 1)(\kappa + \alpha)/(\kappa - 1)$ is again an inverse function of κ. Hence, the total power is minimized by selecting a large scale factor. A more useful parameter to be optimized would be the delay-power product.

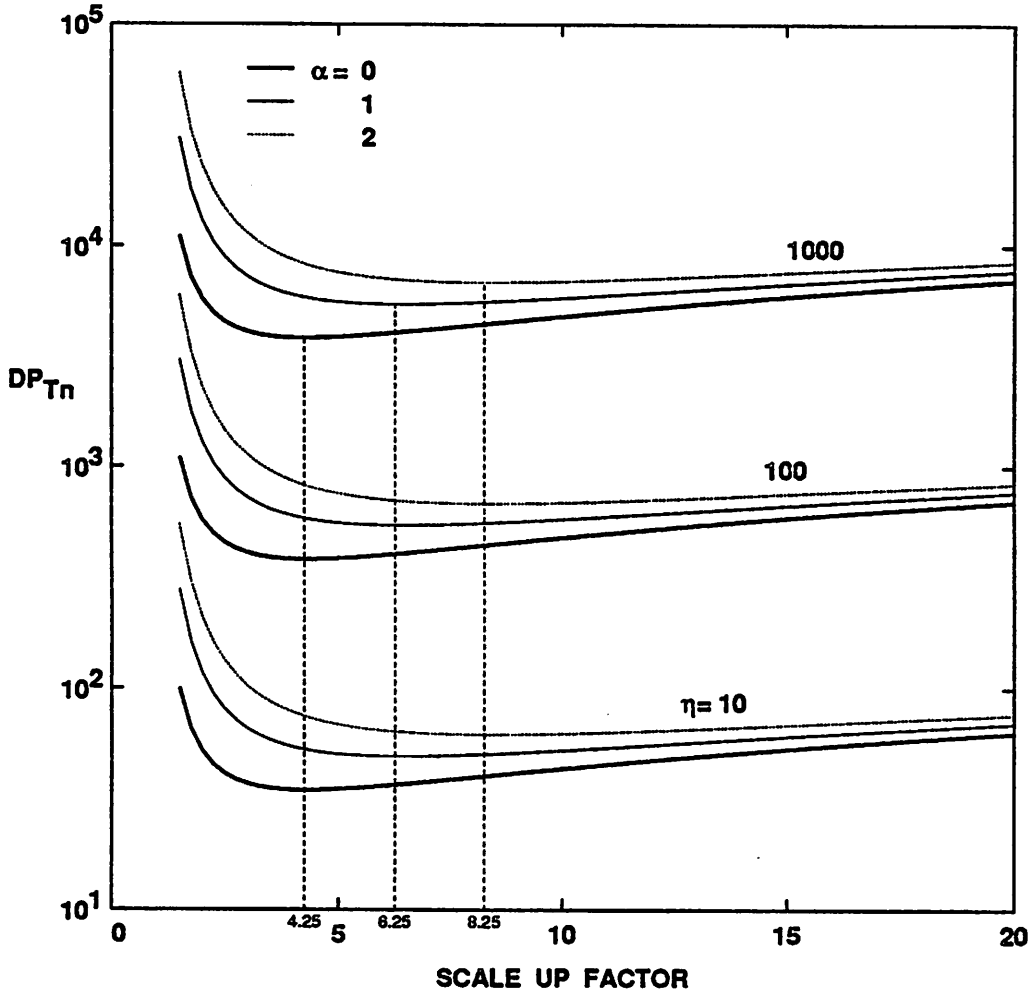


Figure A1.5 Normalized total delay-power product versus scale up factor for different capacitance ratios, η and α.

Figure A1.5 shows the normalized delay-power product,

$$DP_{Tn} = (\eta - 1)\frac{(\kappa + \alpha)^2}{(\kappa - 1)\ln(\kappa)}, \tag{A1.11}$$

for the same values of η and α as in Figure A1.4. Note that a local minimum and the corresponding optimal scale factor exist. The optimal scale factors are larger than the ones in Figure A1.2 and almost independent of η values.

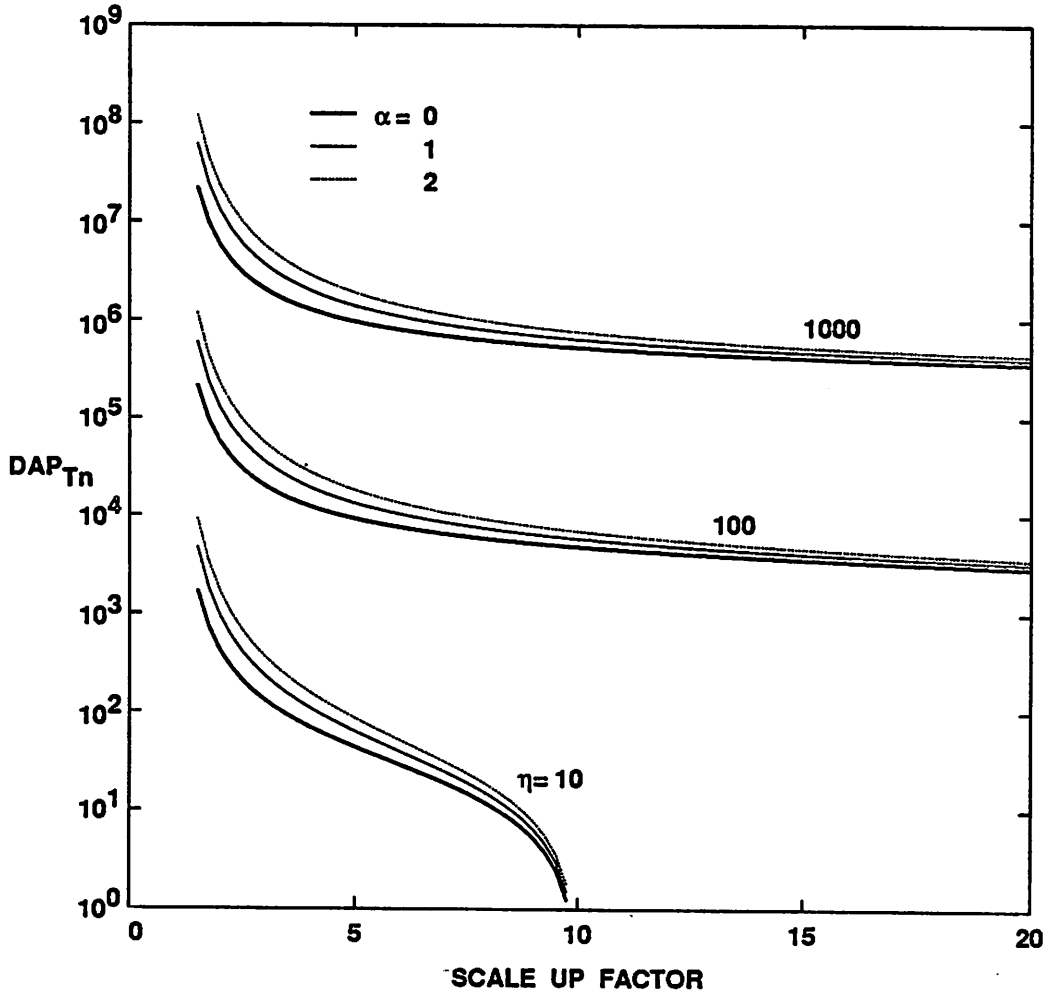


Figure A1.6 Normalized total delay-area-power product versus scale up factor for different capacitance ratios, η and α.

If static power consumption dominates, the normalized total power has a functional dependency equal to that of the total area (Eq. (A1.9)). This means that Figure A1.4 can be utilized for

the delay-static power optimization as well.

Finally, the delay-area-power product would be a good figure of merit that needs optimization. The normalized product when the dynamic power is dominant is expressed as

$$DAP_{T_n} = \frac{(\kappa+\alpha)^2(\eta-\kappa)(\eta-1)}{(\kappa-1)^2\ln(\kappa)} \tag{A1.12}$$

and is shown in Figure A1.6. Clearly, the delay-area-power product is minimized by maximizing the scale up factor under the constraints like the one in Eq. (A1.7).

APPENDIX 2

LOGIC THRESHOLD VOLTAGE

The logic threshold voltage, V_{LT}, is defined as the input or output voltage of an inverter when the input and output nodes have the same voltage. Therefore, an inverter with its input and output tied together naturally generates V_{LT}. In this appendix, the logic threshold voltage is derived for a CMOS inverter that has either long or short channel transistors. Additionally, the analysis is extended for a case when many such (input and output shorted) inverters with slightly different parameters are connected in parallel.

A2.1 CMOS INVERTER

If the simple long channel MOSFET model in saturation region [59] is used while neglecting the drain conductance, V_{LT} is derived from $I_{DSn} = |I_{DSp}|$ with $V_{GSn} = V_{DSn} = V_{LT}$ or

$$\frac{k'_n}{2}(\frac{W}{L})_n(V_{LT} - V_{Tn})^2 = \frac{k'_p}{2}(\frac{W}{L})_p(V_{DD} - V_{LT} - |V_{Tp}|)^2 \qquad \text{(A2.1)}$$

as

$$V_{LT} = \frac{\beta_R^{1/2}V_{DD} + V_{Tn} - \beta_R^{1/2}|V_{Tp}|}{1 + \beta_R^{1/2}} \qquad \text{(A2.2)}$$

where

$$\beta_R \equiv \frac{k'_p(\frac{W}{L})_p}{k'_n(\frac{W}{L})_n}. \qquad \text{(A2.3)}$$

For transistors with very short channel lengths, the saturation current is modeled as [60]

$$I_{DS} = k_S(\frac{W}{L})(V_{GS} - V_T) \qquad \text{(A2.4)}$$

with

$$k_S \equiv \frac{\varepsilon_{ox}}{\frac{(1+\delta)}{v_{norm}} + \frac{t_{ox}}{(1+\delta)v_{sat}L}}, \tag{A2.5}$$

where δ is the body effect factor, v_{norm} is the normal velocity for surface scattering, and v_{sat} is the lateral saturation velocity. Thus, the logic threshold becomes

$$V_{LT} = \frac{\beta_R V_{DD} + V_{Tn} - \beta_R |V_{Tp}|}{1+\beta_R} \tag{A2.6}$$

with a new

$$\beta_R \equiv \frac{k_{Sp}(\frac{W}{L})_p}{k_{Sn}(\frac{W}{L})_n}. \tag{A2.7}$$

In both cases, if the transistors are sized so that $\beta_R = 1$, then $V_{LT} = \frac{(V_{DD} + V_{Tn} - |V_{Tp}|)}{2}$.

Figure A2.1 illustrates V_{LT} versus β_R from Eqs. (A2.2) and (A2.6) for different sets of V_{Tn} and V_{Tp} with $V_{DD} = 5$ V. Note that short channel inverters exhibit more V_{LT} variation. Most modern CMOS inverters fall in between the two extremes, i.e., long and short cases.

A2.2 CMOS INVERTERS CONNECTED IN PARALLEL

When identically sized (in nominal sense) CMOS inverters are connected in parallel, the logic threshold is obtained from the relation

$$\sum_i I_{DSn,i} = \sum_i I_{DSp,i}. \tag{A2.8}$$

In case all of the transistor parameters exactly match among inverters, this equation simply results in the same V_{LT} as that derived in the previous section. However, if the parameters mismatch as found in real examples, this would be hard to solve without approximations.

For long channel transistors, Eq. (A2.8) is equivalent to

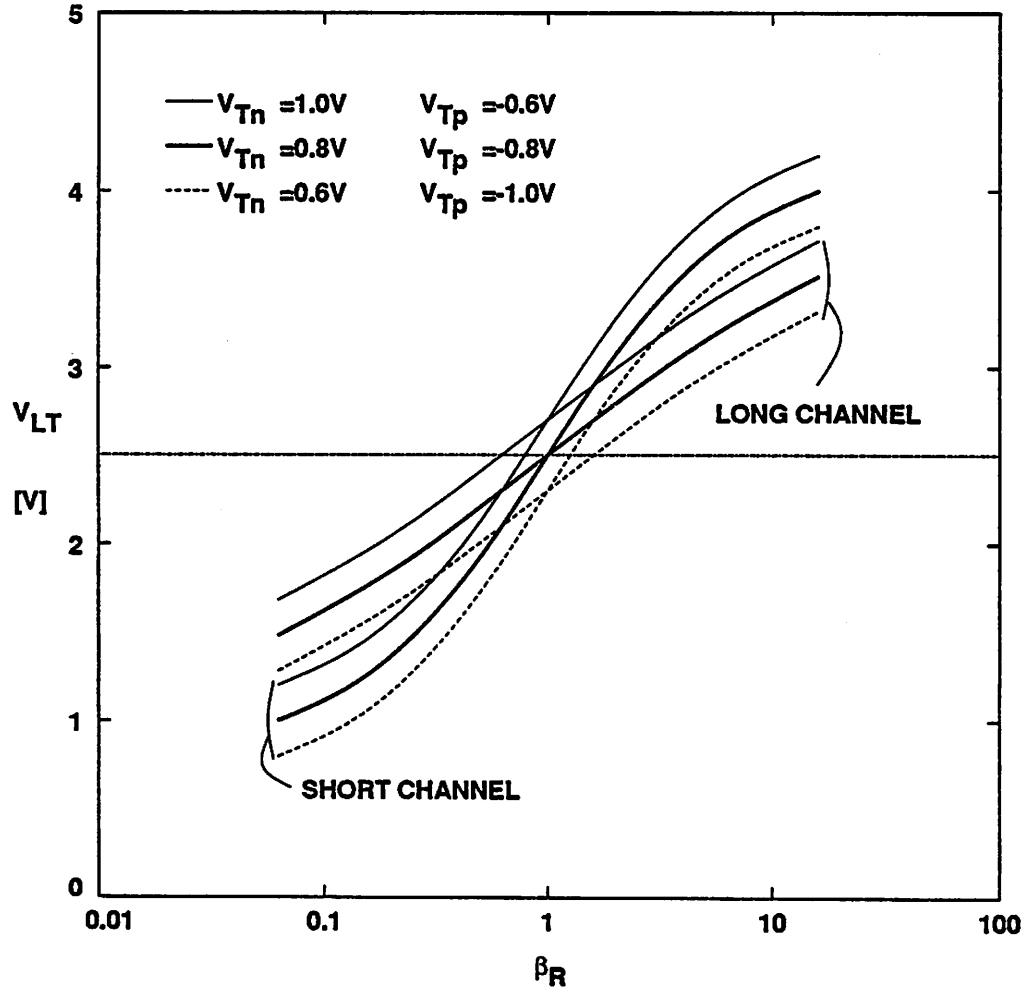


Figure A2.1 Logic threshold voltage versus transistor gain ratio for different sets of MOSFET threshold voltages with $V_{DD} = 5$ V.

$$\sum_i \frac{k'_{n,i}}{2}\left(\frac{W}{L}\right)_{n,i}(V_{LT} - V_{Tn,i})^2 = \sum_i \frac{k'_{p,i}}{2}\left(\frac{W}{L}\right)_{p,i}(V_{DD} - V_{LT} - |V_{Tp,i}|)^2. \tag{A2.9}$$

At first, if $V_{Tn,i}$ and $V_{Tp,i}$ are assumed to be constant, the logic threshold is simply expressed by Eq. (A2.2) with

$$\beta_R \equiv \frac{\sum_i k'_{p,i}\left(\frac{W}{L}\right)_{p,i}}{\sum_i k'_{n,i}\left(\frac{W}{L}\right)_{n,i}} = \frac{[k'_p\left(\frac{W}{L}\right)_p]_{ave}}{[k'_n\left(\frac{W}{L}\right)_n]_{ave}}. \tag{A2.10}$$

where $[k'(\frac{W}{L})]_{ave}$ represents the average of the gain factors, which is close to the nominal value.

This suggests that the nominal logic threshold can be obtained by connecting many inverters, even if each one has slight gain parameter deviations.

Secondly, assuming that $\frac{k'_{n,i}}{2}(\frac{W}{L})_{n,i}$ and $\frac{k'_{p,i}}{2}(\frac{W}{L})_{p,i}$ are constant, $\beta_R = 1$, and variations of $V_{Tn,i}$ and $V_{Tp,i}$ are non-zero but small compared to their means,

$$V_{LT} = \frac{V_{DD} + V_{Tnave} - |V_{Tpave}|}{2}. \tag{A2.11}$$

Therefore, parallel inverters result in the nominal logic threshold, when there are mismatches only in the transistor threshold voltages. In combination, the above results demonstrate the tendency that the nominal logic threshold voltage is achieved when many inverters of minor parameter fluctuation are tied together.

The assumptions given above can be applied for the short channel case as well, and the results would be the same.

REFERENCES

[1] T. Bell, "Technology '85: Communications," *IEEE Spectrum*, vol.22, no.1, pp.56-59, 1985.

[2] F. Guterl and G. Zorpette, "Fiber Optics: Poised to Displace Satellites," *IEEE Spectrum*, vol.22, no.8, pp.30-37, 1985.

[3] E. Rawson and R. Metcalfe, "Fibernet: Multimode Optical Fibers for Local Computer Networks," *IEEE Trans. Communications*, vol.COM-26, no.7, 1978.

[4] R. Schmidt, et al., "Fibernet II: A Fiber Optic Ethernet," *IEEE J. Selected Areas in Communications*, vol.SAC-1, no.5, 1983.

[5] E. Lee and P. Boulton, "The Principles and Performance of Hubnet: A 50 Mbit/s Glass Fiber Local Area Network," *Ibid.*, vol.SAC-1, no.5, 1983.

[6] S. Joshi and V. Iyer, "New Standards for Local Networks Push Upper Limits for Lightwave Data," *Data Communications*, pp.127-136, July 1984.

[7] M. Finley, Jr., "Optical Fibers in Local Area Networks," *IEEE Communications Magazine*, vol.22, no.8, pp.22-35, 1984.

[8] C. Baack, G. Heydt, and J. Vathke, "Digital Integrated Services Broadband Network with HDTV-Capability," *Proceedings of the 11th International Switching Symposium*, 1984.

[9] H. Bauch, K. Euler, and B. Schaffer, "Architectural and Technology Aspects of Broadband Switching," *Ibid.*, 1984.

[10] S. Weinstein, "Personalized Services on the Intelligent, Wideband Network," *Proceedings of the 1986 International Zurich Seminar on Digital Communications*, pp.13-18, 1986.

[11] H. Armbrüster, "Applications of Future Broad-Band Services in the Office and Home," *IEEE J. Selected Areas in Communications*, vol.SAC-4, no.4, pp.429-437, 1986.

[12] L. Hutcheson, P. Haugen, and A. Husain, "Optical Interconnects Replace Hardwire," *IEEE Spectrum*, vol.24, no.3, pp.30-35, 1987.

[13] W. Chou, Ed., *Computer Communications, Volume I - Principles*, Prentice-Hall, Inc., 1983.

[14] C. Tsao, "A Local Area Network Architecture Overview," *IEEE Communications Magazine*, vol.22, no.8, pp.7-11, 1984.

[15] R. Metcalfe and D. Boggs, "Ethernet: Distributed Packet Switching for Local Computer Networks," *Commun. Ass. Comput. Mach.*, vol.19, no.7, pp.395-404, 1976.

[16] J. McDonald, Ed., *Fundamentals of Digital Switching*, New York: Plenum Press, 1983.

[17] M. Liu, "Low-Cost High-Speed General Service Fiber Optic Networks," *U.C. Berkeley ERL, Memorandum No. UCB/ERL M87/26*, 1987.

[18] C. Clos, "A Study of Non-Blocking Switching Networks," *Bell System Tech. J.*, vol.32, pp.406-424, 1953.

[19] W. Young and L. Curtis, "Cascaded Multipole Switches for Single-Mode and Multimode Optical Fibers," *Electronics Letters*, vol.17, pp.571-573, 1981.

[20] M. Shirasaki, et al., "Magnetooptical 2 × 2 Switch for Single-Mode Fibers," *Applied Optics*, vol.23, no.19, pp.3271-3276, 1984.

[21] R. Blackmore, W. Stewart, and I. Bennion, "An Opto-Electronic Exchange of the Future," *Proceedings of the 11th International Switching Symposium*, 1984.

[22] J. Brandon, et al., "Double-Heterostructure GaAs-$Al_xGa_{1-x}As$ Rib Waveguide Directional Coupler Switch," *Proceedings of the Second European Conference on Integrated Optics*, pp.69-71, 1983.

[23] C. von Helmolt, et al., "Bragg Switch / SSB-Modulator with Integrated Optic Single-Mode Waveguides," *Ibid.*, pp.132-135, 1983.

[24] A. Neyer and W. Mevenkamp, "Single-Mode Electrooptic X-switch for Integrated Optic Switching Networks," *Ibid.*, pp.136-139, 1983.

[25] L. Thylen, "High Speed, Wide Bandwidth Digital Switching and Communications Utilizing Guided Wave Optics," *Proceedings of the 1986 International Zurich Seminar on Digital Communications*, pp.113-116, 1986.

[26] K. Pickar and J. Meindl, Ed., Special Issue on Integrated Circuit Technologies of the Future, *Proceedings of the IEEE*, vol.74, no.12, 1986.

[27] D. Barbe, Ed., *Very Large Scale Integration (VLSI): Fundamentals and Applications*, Springer-Verlag Berlin Heidelberg New York, Ch.5, 1980.

[28] M. Cooke, G. Sumerling, T. Muoi, and A. Carter, "Integrated Circuits for a 200-Mbit/s Fiber-Optic Link," *IEEE J. Solid-State Circuits*, vol.SC-21, no.6, pp.909-915, 1986.

[29] D. Clawin, U. Langmann, and H. Schreiber, "5-Gbit/s Si Integrated Regenerative Demultiplexer and Decision Circuit," *IEEE J. Solid-State Circuits*, vol.SC-22, no.3, pp.385-389, 1987.

[30] R. Reimann and H. Rein, "Bipolar High-Gain Limiting Amplifier IC for Optical-Fiber Receivers Operating up to 4 Gbit/s," *IEEE J. Solid-State Circuits*, vol.SC-22, no.4, pp.504-511, 1987.

[31] R. Swartz, et al., "A 2Gb/s Silicon NMOS Laser Driver," *ISSCC DIGEST OF TECHNICAL PAPERS*, pp.64-65, 1986.

[32] R. Jindal, "Gigahertz-Band High-Gain Low-Noise AGC Amplifiers in Fine-Line NMOS," *IEEE J. Solid-State Circuits*, vol.SC-22, no.4, pp.512-521, 1987.

[33] J. Steininger and E. Swanson, "A 50Mb/s CMOS Optical Data Link Receiver Integrated Circuit," *ISSCC DIGEST OF TECHNICAL PAPERS*, pp.60-61, 1986.

[34] A. Fisher and N. Linde, "A 50-Mbit/s CMOS Optical Transmitter Integrated Circuit," *IEEE*

J. Solid-State Circuits, vol.SC-21, no.6, pp.901-908, 1986.

[35] T. Wakimoto, Y. Akazawa, and K. Kawarada, "4-GHz Band GaAs Monolithic Limiting Amplifier," *IEEE J. Solid-State Circuits*, vol.SC-21, no.6, pp.1103-1108, 1986.

[36] C. Duret, F. Pirio, J. Majos, and A. Chemarin, "Integrated Crosspoint Technology for Wide Band Space Division Switching Networks," *Proceedings of the 11th International Switching Symposium*, 1984.

[37] D. Vlack and H. Lehman, "An Experimental Digital Video Switching Architecture," *Ibid.*, 1984.

[38] G. Hayward, A. Gottlieb, D. Boyer, and J. Berthold, "High-Speed 16x16 CMOS Crosspoint Switch," *Electronics Letters*, vol.21, no.20, pp.923-925, 1985.

[39] D. Böttle and M. Klein, "High-Speed (140Mbit/s) Switching Techniques for Broadband Communications," *Proceedings of the 1986 International Zurich Seminar on Digital Communications*, pp.97-100, 1986.

[40] H. Reiner, "Integrated Circuits for Broad-Band Communications Systems," *IEEE J. Selected Areas in Communications*, vol.SAC-4, no.4, pp.480-487, 1986.

[41] TriQuint Semiconductor, *private communication.*

[42] W. Eckton, "Packaging Issues for Integrated Circuits," *Proceedings of the 1986 Bipolar Circuits and Technology Meeting*, 1986.

[43] KYOCERA Corporation, *private communication.*

[44] D. Hodges and H. Jackson, *Analysis and Design of Digital Integrated Circuits*, McGraw-Hill, Inc., 1983.

[45] P. Totta and R. Sopher, "SLT Device Metallurgy and its Monolithic Extension," *IBM J. Res. Dev.*, vol.13, p.226, 1969.

[46] P. Gray and R. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons, Inc., 1984.

[47] Y. El-Mansy, "MOS Device and Technology Constraints in VLSI," *IEEE Trans. Elec. Dev.*, vol.ED-29, no.4, pp.567-573, 1982.

[48] R. Troutman, *Latchup in CMOS Technology: The Problem and Its Cure*, Boston, MA: Kluwer Academic Publishers, 1986.

[49] G. Hu and R. Bruce, "A CMOS Structure with High Latchup Holding Voltage," *IEEE Elec. Dev. Letters*, vol.EDL-5, no.6, pp.211-214, 1984.

[50] R. Rung, C. Dell'Oca, and L. Walker, "A Retrograde P-well for Higher Density CMOS," *IEEE Trans. Elec. Dev.*, vol.ED-28, no.12, pp.1115- 1119, 1981.

[51] R. Rung, H. Momose, and Y. Nagakubo, "Deep Trench Isolated CMOS Devices," *1982 IEDM Tech. Digest*, pp.237-240, 1982.

[52] L. Nagel, "SPICE2, A Computer Program to Simulate Semiconductor Circuits," *ERL Memorandum ERL-M520*, University of California, Berkeley, 1975.

[53] B. Sheu, D. Scharfetter, and P. Ko, "SPICE2 Implementation of BSIM," *U.C. Berkeley ERL, Memorandum No. UCB/ERL M85/42*, 1985.

[54] W. Scott, et al., Ed., "1986 VLSI Tools; Still More Works by the Original Artists," *U.C. Berkeley Computer Science Division, Technical Report No. UCB/CSD 86/272*, 1986.

[55] J. White and A. Sangiovanni-Vincentelli, "RELAX2.1: A Waveform Relaxation Based Circuit Simulation Program," *Proc. Custom Integrated Circuit Conf.*, 1984.

[56] D. Millicker, Bell Communications Research, *private communication.*

[57] A. M. Mohsen and C. A. Mead, "Delay-Time Optimization for Driving and Sensing of Signals on High-Capacitance Paths of VLSI Systems," *IEEE J. Solid-State Circuits*, vol.SC-14, pp.462-470, 1979.

[58] L. A. Glasser and L. P. J. Hoyte, "Delay and Power Optimization in VLSI Circuits," *21st Design Automation Conf.*, pp.529-535, 1984.

[59] H. Schichman and D. Hodges, "Modeling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits," *IEEE J. Solid-State Circuits*, vol.SC-3, no.3, pp.285-289, 1968.

[60] S. Garverick and C. Sodini, "A Simple Model for Scaled MOS Transistors that Includes Field-Dependent Mobility," *IEEE J. Solid-State Circuits*, vol.SC-22, no.1, pp.111-114, 1987.