# Current Mode Interconnect

Marshnil Dave, Maryam Shojaei Baghini, Dinesh Sharma

Department Of Electrical Engineering
Indian Institute Of Technology, Bombay

December 2, 2010

# Contents

# Chapter 1

# Introduction

## 1.1 Scaling

VLSI technology has used device scaling to continually improve the performace of circuits. In constant field scaling, all device dimensions as well as all voltages are scaled down by some factor S. This leads to improved packing density: ($\uparrow S^2$), improved speed (delay $\downarrow S$), and improved power consumption ($\downarrow S^2$). However these improvements apply only to active circuits. What about passive components?

### 1.1.1 Unscaled Interconnect Delay

Consider an interconnect in a chip. This is made of a metal layer of thickness $t_m$ running over an insulator of thickness $t_i$.

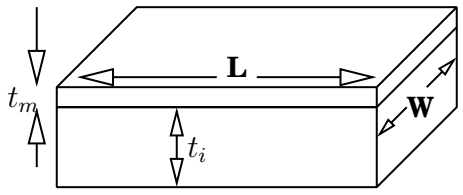



Figure 1.1: Delay through an Interconnect

$$R = \rho\frac{L}{Wt_m}, \qquad C = \epsilon\frac{LW}{t_i}$$

$$\text{Charge Time } \approx RC = \rho\epsilon\frac{L^2}{t_m t_i} \tag{1.1}$$

To first order, delay is independent of W. This is because increasing W reduces resistance but increases capacitance in the same ratio. Unfortunately W is the only parameter that the circuit designer can decide! (L is fixed by the distance between the points to be connected,

$\rho, \epsilon$, $t_m$ and $t_i$ are decided by the technology).

If we see the distribution of wirelengths on a design, there are a large number of wires with short lenths which connect a gate to the other locally. At the same time, there is a con-

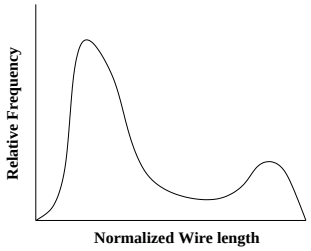



Figure 1.2: Notional distribution of wire lengths on a chip

siderable number of much longer wires which run over the entire chip. These include clocks, power on reset signals, power supply lines, data buses etc. These are the global interconnects.

While local interconnects scale with device size, global interconnects scale with die size. From eqn 1.1

$$\text{Interconnect Delay} = \frac{\rho\epsilon}{t_m t_i} L^2 \equiv AL^2 \tag{1.2}$$

For local interconnects, L scales the same way as $t_m$ and $t_i$, so delay is invariant. However, even as the transistor sizes are scaled down as the technology advances, average chip sizes show an *increasing* trend. This is because the complexity of systems that we put on integrates circuits has increased at a rate higher than the rate at which device geometries shrink. Therefore, for Global Interconnects, L goes *up* with die size, while $t_m$ and $t_i$ scale down. This leads to a sharp increase in delay.

## 1.2 Buffer Insertion for Delay Reduction

Global Interconnect delay can be the determining factor for the speed of an integrated system. The $L^2$ dependence of interconnect delay is a source of particular concern. This problem can be somewhat mitigated by buffer insertion in long wires. We define some critical wire length $L'$ and when a wire segment exceeds this length, we insert a buffer.

### 1.2.1 Optimum Buffer Insertion

What is the optimum wire length after which we should insert a buffer? Consider a long wire in which we insert buffers after every segment of length L'. From eqn 1.2,

$$\text{Segment wire Delay} = \rho\epsilon \frac{L'^2}{t_m t_i} = AL'^2$$

Let buffer delay = $\tau$. For n segments, there will be n-1 buffers, and L = nL' . If the total

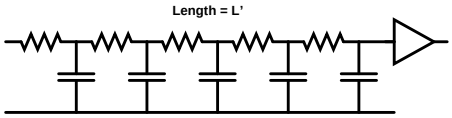



Figure 1.3: A buffered interconnect line

delay is denoted by $\Delta$

$$\Delta = nAL'^2 + (n-1)\tau = \frac{L}{L'}AL'^2 + (\frac{L}{L'} - 1)\tau = ALL' + (\frac{L}{L'} - 1)\tau$$

Putting the derivative with respect to L' = 0 for optimization,

$$AL - \frac{L}{L'^2}\tau = 0, \text{ so } AL'^2 = \tau \tag{1.3}$$

Since $AL'^2$ is the wire delay for the segment, this equation tells us that L' should be so chosen that the wire segment delay = $\tau$. Total delay is proportional to n and so, is linear in L.

## 1.3 Concerns with Voltage mode Buffer Insertion Technique

Currently, buffer insertion is the most widely used method to control interconnect delay. However, there are several difficulties with buffer insertion. Buffers consume power and silicon area. Also, we normally do floor planning and layout first and then put in the interconnects. When the wire length reaches L', we need to put in a buffer. However, it is quite possible that at this point, there is active circuitry underneath, and there is no room to put in a buffer! Then we either have to live with buffer insertion at non-optimal wire lengths or create space by pushing out existing cells and modifying the lay out.

### 1.3.1 Timing closure

Global interconnects are placed *after* active circuit design and layout is complete. One has to anticipate the wire length, and then design the active circuits to meet total delay specifications. If the actual wire length is different from what was anticipated, one has to re-design the active circuits after layout. After a fresh layout, wire lengths and hence, delays are changed. This leads to a design-layout-redesign iteration known as Timing Closure. This iteration becomes longer and longer when total delays are dominated by interconnect delay.

### 1.3.2 Problem with bi-directional data transmission

Global interconnects often include data busses, which may require bidirectional data transmission. (For example, a bus connecting a processor and memory). However, buffer insertion fixes the direction of data flow! Therefore, if we need bidirectional transmission, we need to replace buffers with bidirectional transceivers. These require a direction signal, which will enable the buffers pointing in the desired direction. This direction signal must also be routed with the bus (and should have its own buffers) and it should reach the bidirectional buffers ahead of the data.

### 1.3.3 Signal Integrity

As interconnect wire separation is reduced, there is a serious signal integrity problem because of electrostatic coupling between long wires. Inter-signal interference can lead to unpredictable delay variations. Grounded shielding wires must often be inserted to avoid interference. This leads to extra capacitance and $CV^2f$ power loss.

## 1.4 Current signaling

Because of these problems with voltage mode signaling, we propose that 1's and 0's be signaled by the presence or absence of a current and not by a high or a low voltage. This has several advantages:

- Current rise time is limited by inductance rather than capacitance. Typically, inductive effects are much smaller than capacitive effects. (After all, $\epsilon \simeq 4, \mu = 1$ for insulators used in IC's). So electromagnetic coupling is lower than electrostatic coupling.
- Signal voltage swings are limited by scaled down supply voltages: this does not restrict current swings.
- In fact, we can use multiple current values to send more than one bit down the same wire!

If we hold the Voltage on the interconnect nearly constant dynamic power will be negligible and latency will be much lower.

We also have the option of using multiple current levels to transmit multiple bits simultaneously. This can give higher Throughput and lower interconnect area.

Current mode transmission offers the possibility for improving Latency, Throughput and Power simultaneously!

Since $\Delta V \to 0$, while $\Delta I \neq 0$, $\Rightarrow$ We need a low (near 0) input impedance receiver.

### 1.4.1 Zero input impedance circuit

Low $r_{in}$ amps are used for photo-detectors [**?**]. Once such configuration is shown below: This

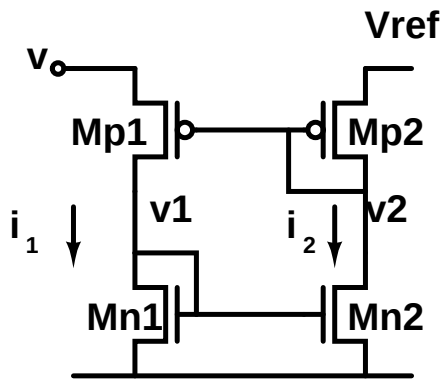



Figure 1.4: Low input impedance Beta Multiplier Circuit

circuit uses complementary current mirrors feeding each other. This configuration is also known as a beta multiplier. To derive its input impedance, we can write small signal currents and voltages as:

$$\begin{array}{rclcl} i_1 & = & g_{mn1}v_1 & = & g_{mp1}(v - v_2) \\ i_2 & = & g_{mn2}v_1 & = & -g_{mp2}v_2 \\ v_2 & = & -\frac{g_{mn2}}{g_{mp2}}v_1 & = & -\frac{g_{mn2}}{g_{mp2}}\frac{i_1}{g_{mn1}} \end{array}$$

$$i_1 = g_{mp1}v + \frac{g_{mn2}/g_{mn1}}{g_{mp2}/g_{mp1}}i_1$$

$$\text{We define } \Gamma \equiv \frac{g_{mn2}/g_{mn1}}{g_{mp2}/g_{mp1}} \tag{1.4}$$

$$\text{then, } i_1(1-\Gamma) = g_{mp1}v$$

$$\text{Which gives } r_{in} = (1-\Gamma)/g_{mp1} \tag{1.5}$$

By making $\Gamma$ close to 1, we can reduce the input impedance to 0. In fact we can set the input impedance to any value, (for example, the characteristic impedance of a transmission line) by a proper choice of $\Gamma$ and $g_{mp1}$. However, we should make sure that $\Gamma$ does not exceed 1, because that will lead to a negative input impedance, and instability. Therefore it is of some interest to determine how accurately we may set the value of $\Gamma$ inspite of power supply, process and temperature variations.

**Robustness of design**

In saturation,

$$I_d = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_g - V_T)^2$$

$$\text{So, } g_m = \mu C_{ox}\frac{W}{L}(V_g - V_T) = \sqrt{2\mu C_{ox}\frac{W}{L}I_d}$$

$$g_{mn2}/g_{mn1} = \sqrt{\frac{(W/L)_{n2}}{(W/L)_{n1}}\frac{I_2}{I_1}}$$

$$g_{mp2}/g_{mp1} = \sqrt{\frac{(W/L)_{p2}}{(W/L)_{p1}}\frac{I_2}{I_1}}$$

$$\text{Therefore } \Gamma \equiv \frac{g_{mn2}/g_{mn1}}{g_{mp2}/g_{mp1}} = \sqrt{\frac{(W/L)_{n2}/(W/L)_{n1}}{(W/L)_{p2}/(W/L)_{p1}}} \tag{1.6}$$

This means that $\Gamma$ depends only on transistor geometries and is independent of supply voltage, bias values, transistor parameters or temperature. This enables us to choose a value of $\Gamma$ very close to 1, which in turn can provide very low input impedence.

**Receiver Design - Input stage**

Just by adding another current mirror transistor and a current to voltage converter, we can use the beta multiplier as a receiver for current mode data signaling.

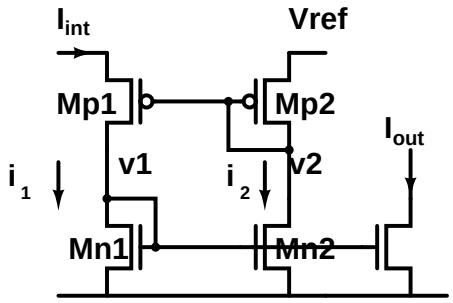



Figure 1.5: A Beta Multiplier based Current Mode Receiver

The input resistance is controlled largely by the geometry of transistors. The beta multiplier also has the property that it drives its own input through a low output impedance to bring it to the same voltage as $V_{ref}$. Thus the interconnect voltage is held fixed. The Input resistance is largely insensitive to process variations. The only dependence comes through

$g_{mp1}$, but since it is multiplied by $1 - \Gamma$ which is close to 0, the sensitivity to variations is quite low.

## 1.5 Other low impedance line terminations

The beta multiplier is not the only choice for providing low input impedance. Simpler circuits like a diode connected MOS transistor are often used. Another option is to use an inverter with its output shorted to its input as the termination. This is equivalent to terminating the line to ground through a diode connected n channel transistor and to $V_{dd}$ through a diode connected p channel transistor. The effective terminating admittance is the sum of $g_m$ values of n and p channel transistors.

Indeed in our later work, we have preferred a reference inverter with its output shorted to input as the line termination. Low input impedance can be achieved by adjusting the

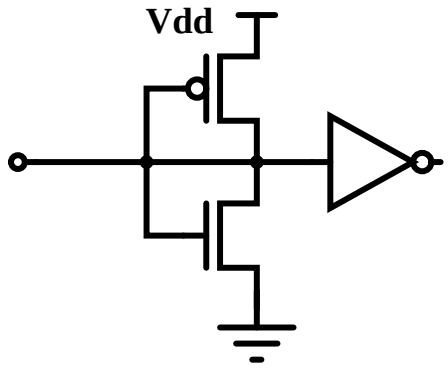



Figure 1.6: Alternative circuit for Low impedance Termination

geometry of the p and n channel transistors. This termination is faster because of the absence of parasitic capacitances contributed by the beta multiplier transistors. The termination holds the line at a DC potential which is matched to the transition voltage of the amplifier inverter which follows the termination.

### 1.5.1 Digital Designers need not panic!

We suggest that only the interface works in current mode. Rest of the circuit remains traditional.

A library circuit will do the voltage mode to current conversion (transmitter) and another will convert the current back to voltage mode (receiver).

To put this plan into action, we need a receiver with very low input impedance. (If inductive effects are to be taken into account, we would like to terminate the line into its characteristic impedance.)

## 1.6 Reduced swing signaling

The main advantage of the current mode signaling comes from the fact that the line voltage is held nearly constant. This is somewhat similar to low swing signaling in voltage mode. Low swing signaling in voltage mode involves driving high capacitive loads like interconnects

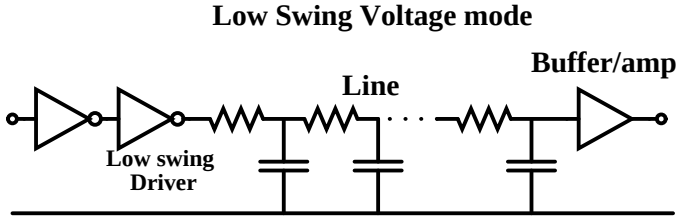



Figure 1.7: Reduced Swing Voltage Mode Signaling

to re-defined levels for 0 and 1 which drastically reduce the voltage swing on the load. The levels are restored to the usual CMOS levels at the receiver end by amplification. This can drastically reduce the power required by line drivers

It is important to distinguish between reduced swing voltage mode signaling and current mode signaling.

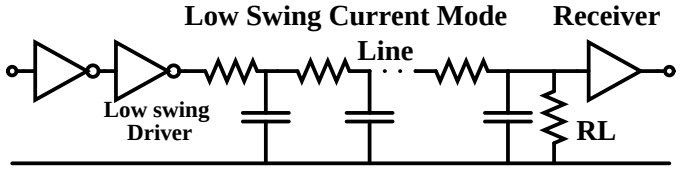



Figure 1.8: Current Mode signaling

- In reduced swing voltage mode signaling, the line is not terminated in a low impedance.
- Current mode signaling terminates the line in a low impedance.
- This reduces the time constant, increases bandwidth.
- However, this also leads to static power consumption.

## 1.7 Improvment in Current Mode Signaling

Traditional current mode signaling consumes Static Power and presents a trade-off between speed, static power and signal to noise ratio. Its performance can be improved by two techniques:

- Inductive Peaking
- Dynamic Over-driving

### 1.7.1 Inductive Peaking

On-chip interconnects can be modeled as distributed RC lines which is essentially a low pass filter. This results in severe attenuation of high frequency components of the signal arriving at the receiver end. This can be corrected by bandwidth enhancement techniques used in RF amplifiers. This involves inductive peaking where the line termination circuit exhibits inductive input impedance. Current flowing through the inductor will produce a voltage $(j\omega L)i$, which increases with frequency. Thus, this can counteract the high frequency attenuation due to the line.

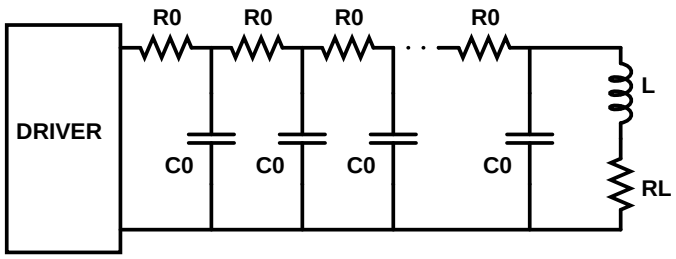
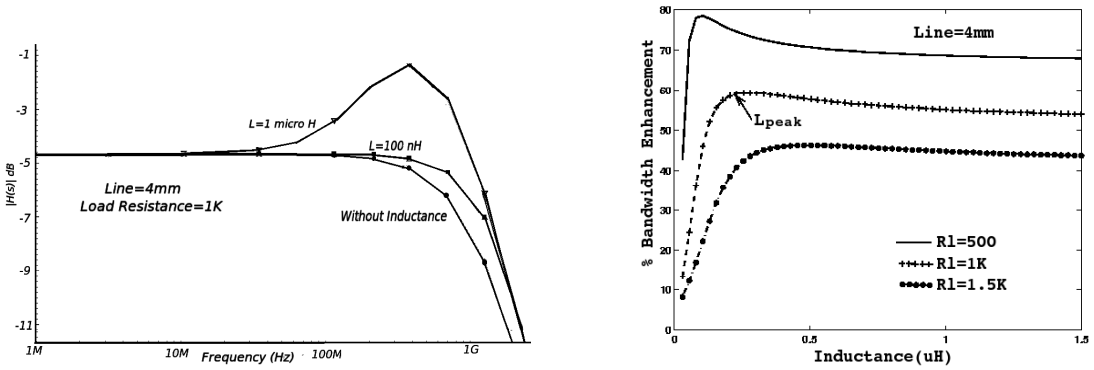



Figure 1.9: Inductively Terminated Line

We performed simulations in which the interconnect line was represented by a realistic LCR segmented line. This was then terminated with resistive/inductive loads of different values. Results of the simulation are shown in fig. 1.10 for a 4mm long line terminated with a 1K resistor in series with different inductance values. The transfer function of the terminated line is plotted as a function of frequency on a log-log scale in fig. 1.10 (a). For a given line length, the amount of bandwidth enhancement is a function of inductance and load resistance. The bandwidth increases with inductance upto a point and after that it remains fixed at that value. As can be seen, we can achieve enhancement of about 500MHz in 3dB bandwidth in this example for an inductive termination of 100 nH. (Because of the log scale, the separation between the curves does not truely reflect the amount by which the bandwidth has been increased). The bandwidth enhancement remains at roughly the same value for larger inductances. We designate the inductance at which the improvement in bandwidth



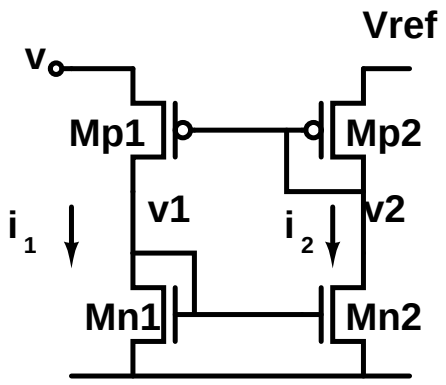

(a) (b)

Figure 1.10: Effect of Inductive Termination on Bandwidth

saturates as $L_{peak}$. As seen from fig 1.10 (b), The dependence on L is not very critical as long as the value is greater than $L_{peak}$. The required inductance for significant enhancement in bandwidth $L_{peak}$ is of the order of a few hundreds of nano Henries. This cannot be conveniently made from spiral inductors etc. Therefore for a practical implementation, we need an active inductor.

### Beta Multiplier: A Gyrator

The beta multiplier circuit suggested earlier for achieving low input resistance values can infact be used to simulate inductances of required values. The Beta Multiplier essentially




forms a gyrator circuit with two Gm elements connected back to back along with the parasitic capacitance of the transistors. So Beta Multiplier Circuits can exhibit inductive input impedance for some frequency range if designed properly.

**Beta Multiplier: Input Impedance**

The input impedance of the beta multiplier is calculated by taking parasitic capacitances into account.

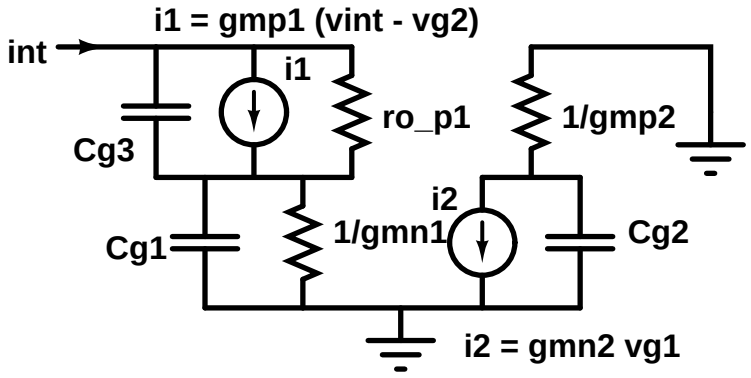



Figure 1.11: Small Signal Equivalent Circuit of Beta Multiplier

We define:

$$\begin{array}{lll} \tau_1 \equiv \frac{C_{g1}}{g_{mn1}} & \tau_2 \equiv \frac{C_{g2}}{g_{mp2}} & R_1 \equiv \frac{1}{g_{mn1}} \\ \tau_3 \equiv C_{g3} r_{op1} & \tau_4 \equiv \frac{C_{g3}}{g_{mp1}} & R_3 \equiv r_{op1} \\ \gamma \equiv \frac{g_{mp1}/g_{mp2}}{g_{mn1}/g_{mn2}} & & k \equiv \frac{R_1}{R_3} \end{array}$$

Then the input impedance can be shown to be:

$$Z_{in} = \frac{\{(\tau_1\tau_2 + k\tau_2\tau_3)s^2 + (\tau_1 + \tau_2 + k(\tau_3 + \tau_2))s + 1 + k - \gamma\}}{\{(g_{mp1} + \frac{1}{R_3})\{(1 + \tau_1 s)(1 + \tau_2 s)(1 + \tau_4 s)\}\}} \tag{1.7}$$

Correspondingly, the resistive part of the input impedance can be expressed as:

$$R_{in} = \frac{(1 - \gamma) + \frac{1}{g_{mn1} r_{op1}}}{g_{mp1} + \frac{1}{r_{op1}}}$$

**Beta Multiplier : Equivalent Circuit**

The nature of input impedance (inductive of capacitive) is determined by the relative location of poles and zeros. If the first zero occurs at least a decade prior to the first pole, the input impedance is inductive. To ensure that a zero occurs a decade prior to the first pole, we have to choose operating currents etc., such that $\gamma - \frac{1}{g_{mn1} r_{op1}} > 0.9$ and any two time constants are equal. Under these conditions, we may approximate the input impedance of the beta

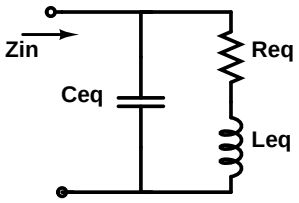



Figure 1.12: Equivalent circuit for the Beta Multiplier

multiplier by the equivalent circuit shown in fig 1.12

Where

$$L_{eff} = \frac{r_{op1}}{g_{mp1}r_{op1}+1}\left\{\frac{C_{g1}}{g_{mn1}}+\frac{C_{g2}}{g_{mp2}}\right. \tag{1.8}$$

$$\left. + \frac{C_{g2}}{g_{mp2}g_{mn1}r_{op1}}+\frac{C_{g3}}{g_{mn1}g_{mp1}r_{op1}}\right\} \tag{1.9}$$

$$R_{eff} = \frac{(1-\gamma)+\frac{1}{g_{mn1}r_{op1}}}{g_{mp1}+\frac{1}{r_{op1}}} \tag{1.10}$$

$$C_{eff} = KC_{gx} \tag{1.11}$$

**Beta Multiplier : Input Impedance Control**

We are interested in using an inductor whose value should be in hundreds of nano Henries. We want to find if these values can be achieved under reasonable bias and geometry conditions. We therefore evaluated the input impedance of the beta multiplier under various operating conditions. As can be seen from the figure, the beta multiplier shows an effective inductance

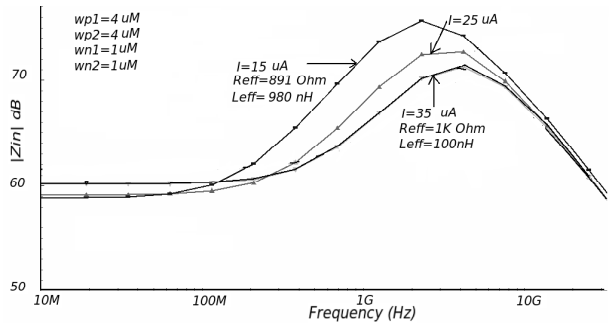



Figure 1.13: Bandwidth enhancement with Beta multiplier termination

of hundreds of nano Henries for a practical range of input current and transistor geometries. Its effective resistance can be controlled by ratios of transconductances while its effective inductance depends on the absolute value of transconductance. It is possible to control $R_{in}$ and $L_{eff}$ with very little interaction between the two. Inductance changes from 100nH to 980nH while the value of effective resistance remains within 12% of its nominal value for $20\mu A$ change in the current.

**Current Mode Receiver Circuit with Beta Multiplier**

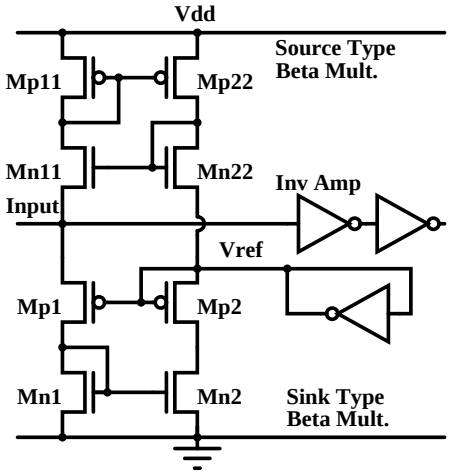



Figure 1.14: Current mode receiver with inductinve peaking using beta multipliers

We can design a current mode receiver with inductive peaking using two beta multipliers as shown in fig. 1.14 above. One of the beta multipliers sources current while the other sinks current. The Effective impedance offered by the receiver is equal to the parallel combination of the impedance offered by individual beta multipliers. Voltage at the input node swings around $V_{ref}$. The small voltage swing on the line is sensed and amplified by the inverting amplifier. $V_{ref}$ is generated by shorting the input and output of an inverter to ensure that the value of $V_{ref}$ is the same as the switching threshold of receiver amplifier across all process corners.

$r_{out}$ of $V_{ref}$ generation circuit comes in series with beta multiplier $Z_{in}$ and hence beta multiplier has to be sized accordingly.
$V_{ref}$ generation circuit consumes static power.

### 1.7.2 Simulation Results

To see the effectiveness of inductive termination, we should compare the power as well as speed of the voltage mode buffer insertion scheme, Diode connected MOS terminated current mode scheme and the beta multiplier based inductive peaking scheme. Simulations were performed for a 6mm long line at a rate of 1 Gbps. Results of the comparison are summarized in the table below:

(line=6 mm, Power measured at 1Gbps)

| Signaling Scheme | Delay (ps) | Throughput (Gbps) | Power ( $\mu W$) | Area ($\mu m^2$) |
|---|---|---|---|---|
| CMS-BMul(30 mV)[1] | 420 | 2.56 | 310 | 2.00 |
| CMS-Diode-CC(30 mV)[2] | 500 | 2.45 | 380 | 2.00 |
| Voltage Mode | 1000 | 2.85 | 3000 | 12.53 |

Inductive termination gives 16% improvement in delay and about 18 % improvement in power compared to Diode termination. Compared to Voltage Mode scheme, we see more than 50 % improvement in delay and an order of magnitude lower power [**?**, **?**].

### 1.7.3 Dynamic Overdriving

Inductive peaking attempts to correct the low pass nature of the line by putting a high pass termination at the receiver end. However, by the time the signal reaches the receiver, its high frequency components have been severely attenuated. Therefore boosting them back to normal level will also boost high frequency noise.

Rather than boosting the high frequency components at the receiver end, why don't we boost them *before* attenuation at the transmitter itself? This technique of boosting the high frequency components before passing them through a low pass channel is know as "pre-emphasis".

#### Concept of Dynamic Overdriving/Pre-emphasis

Current mode transmission can be speeded up by using high drive current. However, this increases static power consumption. One possible solution is to dump high drive current only when the state of the line needs to be changed from 0 to 1 or from 1 to 0. When the line remains at 1 or at 0 from one bit to the next, we use a small drive current to maintain the line at the required voltage. This is called Dynamic Over Driving. Dynamic Overdriving essentially means amplifying high frequency components of the input signal

### Possible implementation of Dynamic Overdriving

The transmitter end contains a weak driver and a strong driver. The strong driver is enabled only when a level change is needed from 0 to 1 or from 1 to 0.

### Weak Driver

The weak driver provides the minimal drive required to keep the line (terminated by low impedance) at the desired voltage level. When the input is 1, the p channel driver gate is low

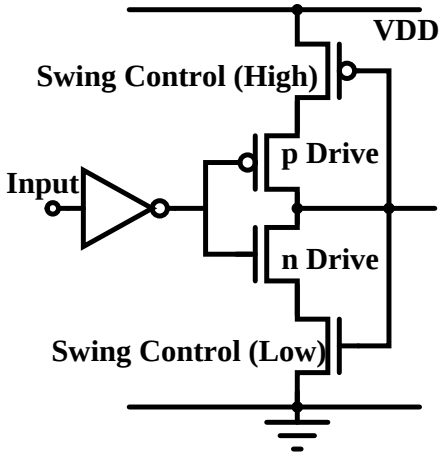



Figure 1.15: Steady State (Weak) Driver

(enabled). This charges up the output. As the line voltage reaches $V_{DD} - V_{Tp}$, the upper p channel transistor turns off, restricting line voltage swing in the up direction.

Similarly when the input is 0 the n channel driver transistor is enabled by a high level at its gate. The transistor discharges the line. However, when the line voltage approaches $V_{Tn}$ during discharge, the lower transistor turns off, stopping the discharging process.

Thus the line can only swing beween $V_{DD} - V_{Tp}$ and $V_{Tn}$. [**?**]

### Strong Driver

The strong driver should be enabled only when the input and the level on the output line do not represent the same logic. The feedback inverter acts as an inverting amplifier converting low swing logic levels on the wire to full swing (inverted) CMOS logic level on its output. The P channel gate is low (enabled) only when both inputs to the NAND are 1. This will happen only when the input is high AND the line is at 0. This is indeed the condition when we want

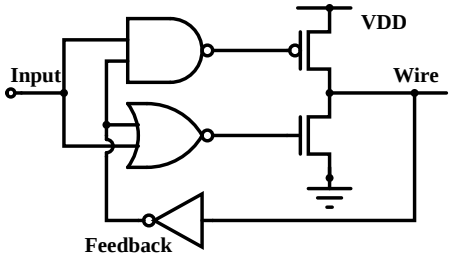



Figure 1.16: Dynamic (Strong) Driver

the strong driver to charge the line.

The N channel gate is high (enabled) only when both inputs to the NOR gate are 0. This will happen only when the input is low AND the line is at 1.

Notice that the input to the feedback inverter is a low swing level around $V_{DD}/2$. Therefore it consumes static power.

The action of the strong driver is self limiting. This is because both NAND and NOR receive the input and the *inverted* logic level of the line. If the input and the logic level of the line are the same, NAND and NOR are fed with input and $\overline{\text{input}}$. Thus one of the inputs to NAND/NOR is 1, while the other is 0. This ensures that the output of NAND is 1, while that of NOR is 0, so that both the p and n channel transistors are OFF. Therefore the strong driver does not need a series transistor as was the case for the weak driver.

When the Input = 1 and Wire voltage $< V_m$,
the inverter output = 1, NAND output = 0 and NOR output = 0.
The P channel driver is ON and dumps current to charge the line.

When the Input = 0 and Wire voltage $> V_m$,
the inverter output = 0, NAND output = 1 and NOR output = 1.
the N channel driver is ON and sinks current to discharge the line.

As soon as low swing logic level on the line becomes equal to the logic level at the input
Inverter output = $\overline{input}$,
and so NAND output = 1, NOR output = 0;
which disables both drive transistors automatically.

**Dynamic Overdriving with Inductive termination?**

Dynamic Overdriving (DOD) and Inductive line termination both essentially amplify high frequency components of input signal. Can we use both?

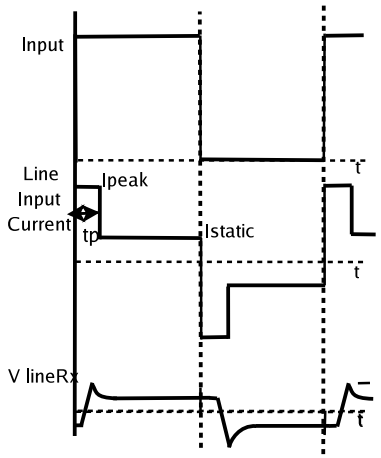



Figure 1.17: Current drive from a Dynamic Over Drive (DOD) type transmitter

To answer this question, the following four current mode signaling schemes were simulated:

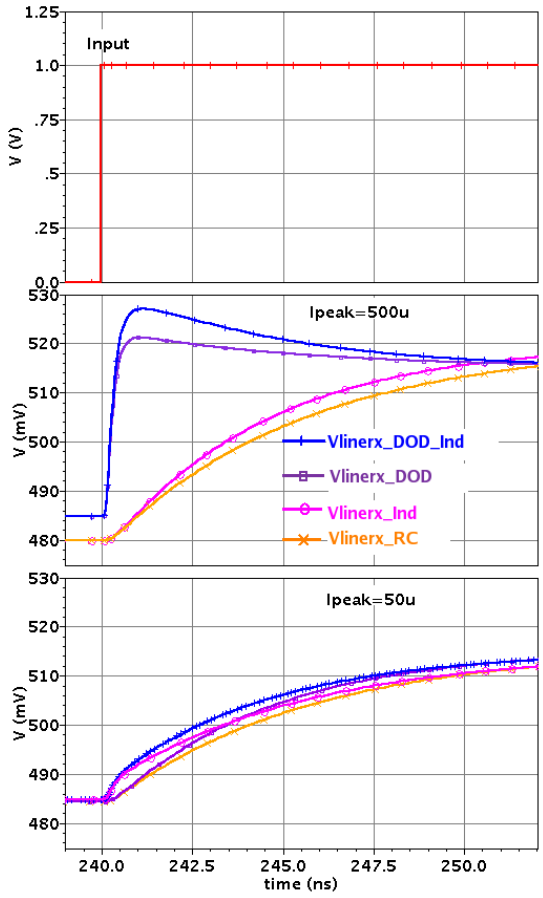
- CMS Scheme with DOD and Resistive Load
- CMS Scheme with Simple Driver and Resistive Load
- CMS Scheme Inductive Load
- CMS Scheme with DOD and Inductive Load

Dynamic Overdriving driver was implemented by an ideal voltage controlled current source (VCCS) with the output current wave shape as shown in fig 1.17. The Simple driver was implemented as a Voltage Controlled Current Sounce with a square output current wave shape. The drive current in this case is $-I_{avg}$ for a 0 at the input and $+I_{avg}$ for a 1 at the input. For a fair comparison, $I_{avg}$ for the simple driver is equal to the weighted mean of the current used for dynamic overdrive transmitter.

$$I_{avg} = \frac{I_{peak}t_p + I_{static}(t - t_p)}{t} \tag{1.12}$$

For this comparison, we used terminations of

$$R_L = 4k\Omega, L = 4\mu H$$




### Comparison of Delay

With Large Overdrive ($I_{peak} = 500\mu A$)

- Dynamic overdriving shows 5 $\times$ improvement in delay over RC
- Inductive peaking does not offer substantial additional advantage when combined with dynamic overdriving.
- Inductive peaking alone shows 25% of improvement in delay over RC

With Small Overdrive ($I_{peak} = 50\mu A$)

- Dynamic Overdriving alone and inductive peaking alone give nearly the same delay
- Inductive peaking along with dynamic overdriving shows around 20% improvement in delay over dynamic overdriving alone

**Comparison of Throughput (Eye-opening)**

We apply a random sequence of bits to the input at a given data rate and observe the wave form at the receiver. The wave form, when observed for two clock periods, looks like a pair of eyes and is known as the "eye diagram". Wide open eyes in the vertical direction represent good signal to noise ratio as the '1' level and the '0' level are well separated. Goof eye opening in the time direction represents low timing jitter in the arrival time of bits – which is also a desirable feature.

As the data rate is increased, The eye closes in the vertical direction, as there is not sufficient time for the driver to charge/discharge the line. Assuming that the receiver is capable of resolving a 30mV input to a full rail to rail swing output, we determine the data rate at which the eye opening is reduced to 30mV. This is the maximum throughput which can be supported by the interconnect. Using this criterion, We can now compare the throughput for the different schemes. We find that

- Dynamic overdriving improves throughput by 5 $\times$ over RC
- Inductive peaking does not offer substantial additional advantage when combined with dynamic overdriving.
- Inductive peaking shows throughput enhancement of 26% over RC

**Conclusion: Inductive Peaking vs Dynamic Overdrive**

- For very high data rate applications, dynamic overdriving alone should be employed as inductive peaking does not offer any additional advantages
- For low power and low data rate applications, the use of inductive peaking can give 26% improvement in throughput and 16% improvement in delay over RC.
- For low power and low data rate applications, the use of dynamic overdrive along with inductive peaking can further improve the throughput by 20%

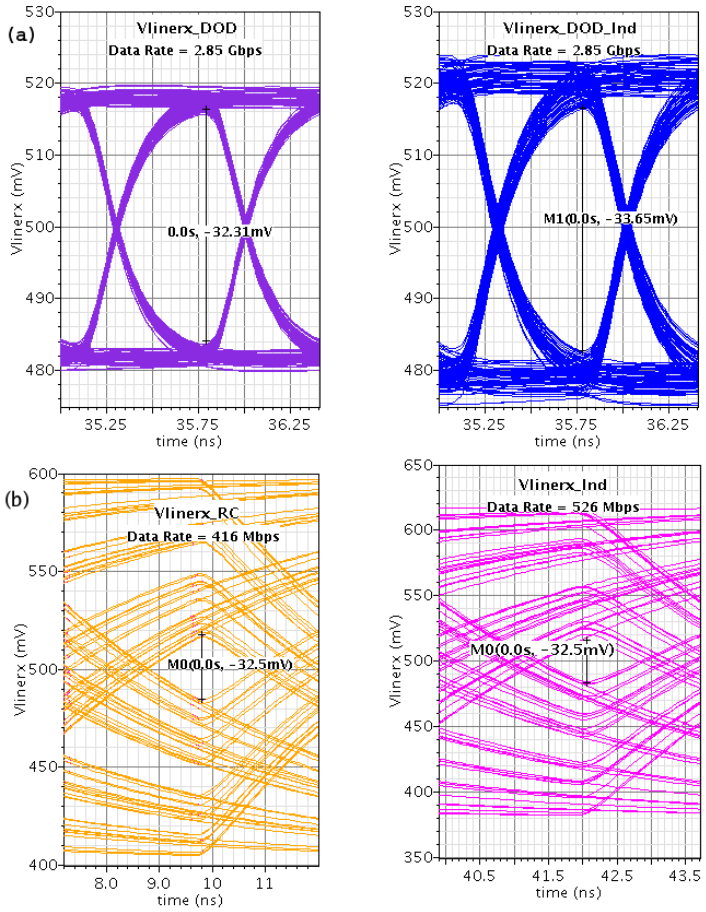



Figure 1.18: Eye diagram for different schemes at data rates where the eye opening is ≈ 32 mV

# Chapter 2

# Variation Tolerant Current Mode Signaling

## 2.1 Need for Process Variation Tolerance

Current mode signaling derives its advantages over voltage mode due to the reduced swing on the line. Careful design is necessary, otherwise small changes in device parameters can have a disproportionate effect on the performance of the system. In modern short channel processes, variations in transistor parameters are large – some of the parameters can vary by as much as 40% of their nominal values. We have to design circuits, so that they are robust with respect to batch-to-batch variations, as well as variations between devices on the same die. Batch-to-batch or inter-die variations can shift operating points and drive strengths, while intra-die variations cause mismatch in parameters of transmitter and receiver transistors.

## 2.2 Robustness requirements

Process, Supply Voltage and Temperature (PVT) variations will affect the core logic as well as data communication circuitry. The requirement for data transmission is therefore not of *complete invariance* with respect to PVT variations. We have to ensure that throughput and delay properties of the interconnect are at least as good as data generation and clock rates. Thus the deterioration in interconnect properties should be *no worse than* the deterioration in general logic.

### 2.2.1 Effect of Process, Voltage and Temperature Variation

Due to process, voltage and temperature variations, the drive capabilities and operating points of various circuits used for data transmission will vary. The cumulative effect of all

these variations on the performance of the interconnect scheme.

### 2.2.2 Effect of common mode voltage mismatch

Because global interconnects, by definition, connect remote points on the die, on chip variations can, in fact, be of even greater concern. On chip variations will result in different common mode voltages at the transmitter and the receiver end. In case of ideal match, small

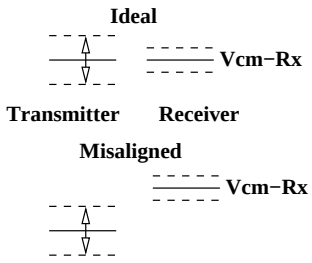
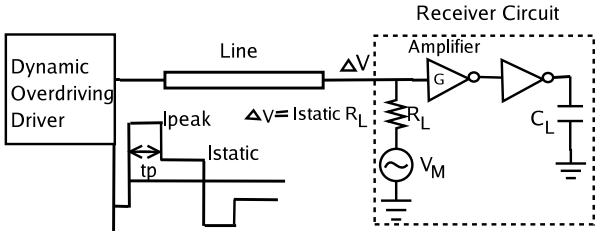



Figure 2.1: Mismatched common mode voltages at Transmitter and Receiver

fluctuations in line voltage are converted to rail to rail swing by the receiver. If, however, the mismatch is large, the small swing on the line may be completely ignored by the receiver. It is important, therefore, that the amount of swing on the line is much more than the mismatch in common mode voltages. But high swing will cause power dissipation. Therefore, it is better to have smart bias circuits, which will reduce mismatch and the need for a large swing.

## 2.3 System parameters affected by variations




Variations in the following parameters have a strong influence on the performance of the signaling scheme:

1. $I_{peak}$: Peak current supplied by the strong driver during input transition
2. $t_p$: Duration for which the strong driver is ON

3. $\Delta V$: Line voltage swing at the receiver end in steady state

4. Mismatch between $V_{CMRx}$ and operating point of an amplifier

## 2.4 A brief review of Current Mode Signaling Schemes

Several current mode signaling schemes have been suggested in the literature. We shall concentrate on three schemes here.

### 2.4.1 CMS Scheme with Feedback (CMS-Fb)

This scheme uses feedback at both the transmitter and the receiver ends to adjust the operating points of these circuits. [**?**] The transmitter used by this scheme is shown below:
The feedback inverter converts low swing logic levels on the line to full rail to rail CMOS

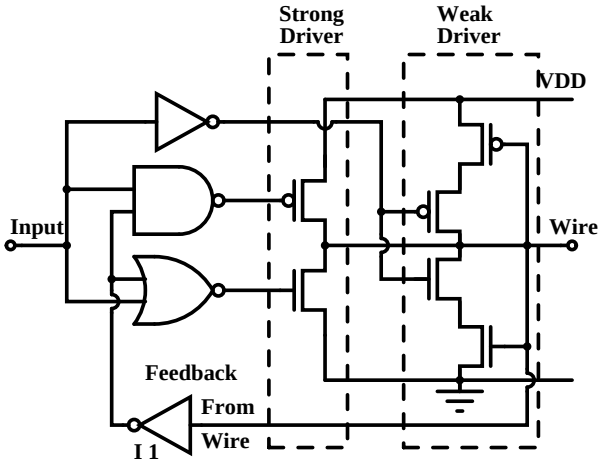



Figure 2.2: Transmitter used by CMS scheme with feedback

levels. The NAND/NOR gates ensure that the strong driver is turned on only during data transitions and is turned off as soon as the line crosses the swithing point of the feedback inverter to make the logic level on the line equal to the input. The weak driver supplies $I_{static}$ and the line voltage swing at the receiver end is $V_{CMRx} \pm I_{static}R_L$ The receiver also uses feedback to adjust its common-mode voltage. Take the case where $V_{CMTx}$ at the transmitter end

## 2.5 Effect of Process Variations on different CMS Schemes

### 2.5.1 CMS Scheme with Feedback (CMS-Fb)

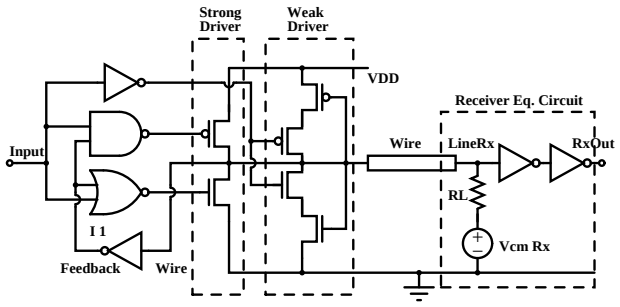



Figure 2.3: Current Mode Scheme with Feedback (CMS-fb)

**Effect of Inter-die Process Variations on CMS with feedback**

- Variations in $I_{peak}$ are well compensated due to the feedback at the driver end.
- If the driver is weaker due to process variations, the feed back system keeps it on for longer till the line reaches the desired voltage.
- This might, however, not be optimum from a power point of view.

**Effect of Intra-die Process Variations on CMS-Fb**

If the $V_{CMTx}$ for the feedback inverter at the transmitter end is not the same as the $V_{CMRx}$ for the receiver amplifier, this scheme does not work very well. Take the case where $V_{CMTx}$

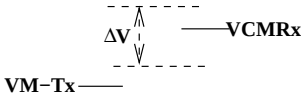
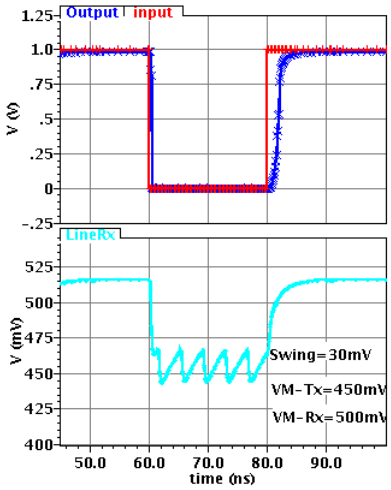



Figure 2.4: Mismatched common mode voltages at Transmitter and Receiver

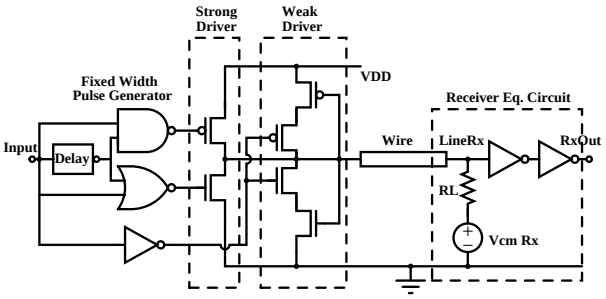
at the transmitter end is lower than the $V_{CMRx}$ at the receiver end. During the low to high transitions the strong driver will be turned off well before the line voltage crosses $V_{CMRx}$. This can result in very slow charging of the line after the strong driver is turned off, leading to a low throughput. In an extreme case, the line voltage may never reach $V_{CMRx}$, leading to

malfunction.




The same phenomenon will occur for the high to low transition if $V_{CMTx} > V_{CMRx}$.

### 2.5.2 CMS Scheme with fixed pulse width (CMS-Fpw)



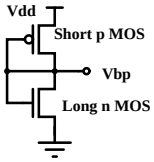

- $t_p$ is given by delay element
- Less sensitive to intra-die variations
- In the skewed corners, sourcing $I_{peak}$ and sinking $I_{peak}$ are different, leading to different rise and fall delay
- Throughput can degrade significantly in skewed corners

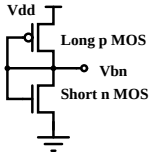
[?]

## 2.6 The Proposed Variation Tolerant CMS Scheme

### Minimizing Process Dependence

To minimize process dependence, we need smart bias circuits which sense the process corner and adjust the bias to compensate for variations.

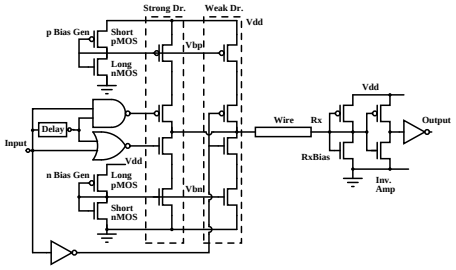



- Long Channel transistors show relatively less variation with process compared to Short Channel transistors in the same process.
- We can make use of this difference to design a bias generator which senses the process corner and tries to increase the transistor current in the slow corners and to decrease it in the fast corners.
- Simple bias generators using inverters with input and output shorted and which use this feature are shown here.

### Proposed CMS Scheme with Smart Bias

We propose a Dynamic Overdrive scheme in which both the strong and the weak drivers use constant current sources controlled by process aware bias generators.




- There is no feedback inverter in the driver circuit
- Bias voltages change in the desired direction to keep the current through weak and strong drivers the same across all corners

### Effect of Process Variation on the Proposed CMS Scheme

- $I_{peak}$ remains nearly the same across all corners. In extreme corners, SS and FF, small change in $I_{peak}$ is compensated by the opposite change in $t_p$.

- $\Delta V = I_{static} R_L$ remains the same across all corners, $R_L = \frac{1}{g_{mn}+g_{mp}}$
- The inverter with input-output shorted and the inverter amplifier are designed using fingers and placed close to each other so that their switching thresholds are closely matched across all corners.
- This makes the proposed circuit less sensitive to intra die process variations as well.

## 2.7 Performance Evaluation

**Simulation Setup**

- Foundry specified four corner model files and mismatch model file for Montecarlo simulations were used.
- All the signaling schemes offer the same input capacitance (equivalent to one minimum sized inverter).
- All signaling scheme drive FO4 load.
- Line RLC used were: $R_{line} = 244\Omega$ /mm, $L_{line} = 1.5nH$/mm, $C_{line} = 201fF$/mm.
- All schemes were designed for a throughput of 2.65Gbps.
- Current mode schemes are designed for $I_{peak} = 500\mu A$

**Effect of Intra-die Process Variations**

Mismatch in $V_m$ of an inverter can be up to 40 mV. [1]. For a mismatch of 40 mV in the $V_m$ value of the inverters,

| CMS system | Percentage Degradation | |
|---|---|---|
| | Delay | Throughput |
| CMS-Fb | **25** | **33** |
| CMS-Fpw | 10 | 14 |
| CMS-Bias | **4** | **9.5** |

[1] Mismatch Data sheet from the foundry

**Effect of Inter-die Process Variations**

| Signaling System/ Logic Circuit | Percentage Degradation | | |
|---|---|---|---|
| | SS | SNFP | FNSP |
| CMS-Fb | 17.5 | 5.7 | 2.9 |
| CMS-Fpw | **32** | **33.6** | **34.9** |
| CMS-Bias | **18.75** | **8.2** | **7.14** |
| Voltage Mode | 27 | $< 1$ | 2.8 |
| Ring Oscillator Freq | **23** | **2.88** | **3** |

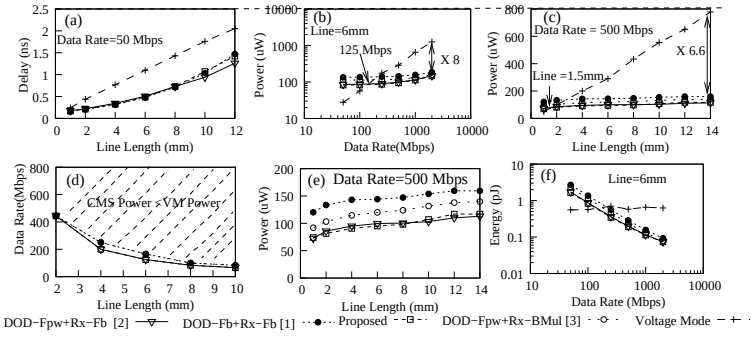
- Interconnects with CMS-Fpw scheme become the bottleneck in overall performance of the chip in skewed corners
- Degradation in the throughput of the proposed scheme in the skewed corners is around 7% which is less than that in CMS-Fpw scheme

**Overall Comparison**

Performance Comparison of four signaling schemes (line=6 mm, Power measured at 1Gbps)

| Signaling Scheme | Delay (ps) | Throughput (Gbps) | Power ( $\mu W$) | Area ($\mu m^2$) |
|---|---|---|---|---|
| CMS-Fb(90 mV) | 700 | 2.56 | 146 | 2.00 |
| CMS-Fpw | 503 | 2.65 | 114 | 2.40 |
| Proposed CMS | 490 | 2.56 | 113 | 3.07 |
| Voltage Mode | 1100 | 2.85 | 655 | 12.53 |

- The CMS-Fb scheme consumes higher power than other schemes due to static power consumption in the feedback inverter
- The proposed scheme shows 78% improvement in area over voltage mode scheme whereas other schemes, CMS-Fb and CMS-Fpw show 84% and 80% respectively



## 2.8 Bidirectional Links

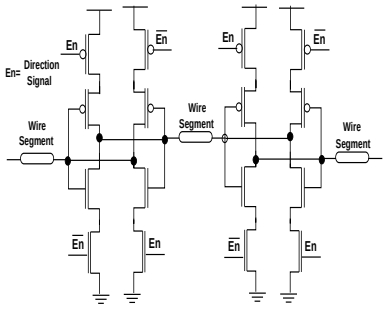
### Bidirectional Links

In many applications, on-chip buses need to carry signal in both directions.

For example, the bus between processor and memory, main processor and floating point multiplier etc.

Often bidirectional buffers with direction control are used for this.

### Limitations of Conventional Bidirectional Buffer




Back-to-Back Connected Tri-state Buffers

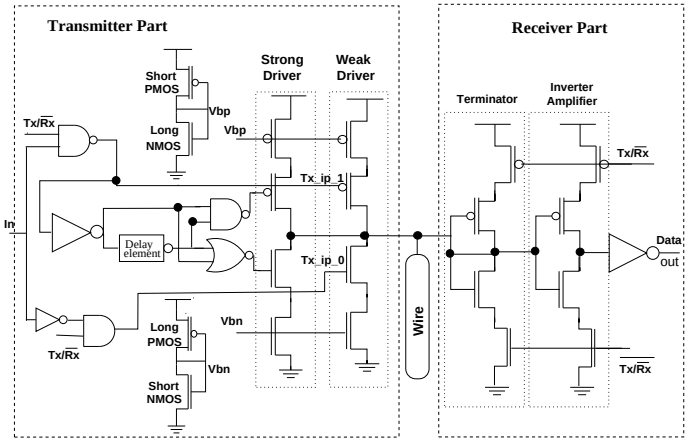
- One of the two tristate buffers is enabled at a given time
- Two transistors in stack ⇒ increased sizes of PMOS and NMOS
- Delay of a bidirectional repeater is more than that of a unidirectional buffer
- Direction control signal is required by each repeater
- Buffers offer huge load to direction control signal
- Buffers carrying direction control signal consume additional power

We need **a repeaterless** Signaling Scheme

### The Proposed Current Mode Bidirectional Link

- Employs only two bidirectional transceivers, one at each end of the line.
- Direction signal is required only at two ends of the line
- The direction control signal can be the same as one of the control signal or derived from it based on communication protocol
- Assumption: Direction signal ($Tx/\overline{Rx}$) is locally available at both ends before data transmission starts

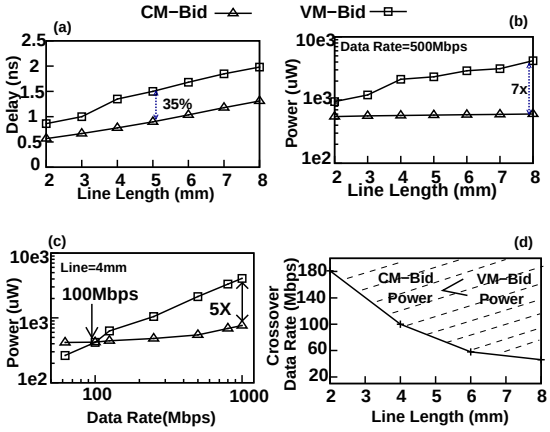
**Proposed Current-Mode Transceiver**




Either the transmitter part or the receiver part is enabled at a time

## 2.8.1 Simulated Performance of Bidirectional Link

**Speed-Power of Proposed Bidirectional CMS Scheme**




Current-Mode Vs. Voltage-Mode

- 35% improvement in delay for nearly all line lengths
- 1.7× lower power for 2mm lines and 7× lower power for 8mm line
- Power crossover frequency 100Mbps for 4mm long lines
- 5 × reduction in power at 1Gbps

- For lines longer than 2mm communicating at data-rates more than 180Mbps, the proposed scheme consumes less power than voltage-mode

Designed in 180nm for $V_{dd}$=1.8V using nominal Vt devices

Line Characteristics: R=211$\Omega/mm$ and C=0.245pF/mm