Current Mode Interconnect

Dinesh Sharma, Marshnil Dave

Department Of Electrical Engineering
Indian Institute Of Technology, Bombay

Sept. 25, 2010

Current Mode Interconnects Group at IIT Bombay

- Prof. Maryam Shojaei Baghini
- Marshnil Dave
- Amit Vishnani
- Navin Kacharappu
- Sandeep Waikar
- Girish Naik
- Dinesh Sharma
- Supreet Joshi
- Rajkumar Satkuri
- Mahavir Jain
- M. Veerraju

Part I

Current Mode Data Communication

Scaling

- Unscaled Interconnect Delay

Solutions for Interconnect Delay problem

- Buffer Insertion
- Current signaling
- Inductive Peaking
- Dynamic Overdriving

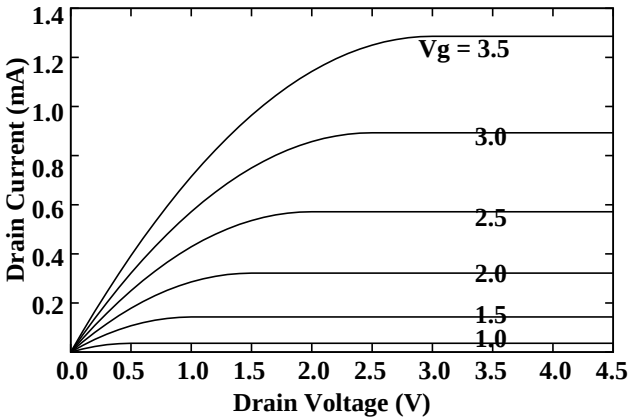
Scaling

- To increase packing density, we would like to reduce the size of transistors and passive components.
- In order to decrease lateral sizes, we have to reduce vertical sizes too.
- If dimensions are scaled down, voltages must also be reduced to avoid breakdown.

This is known as constant field scaling.

So what price do we have to pay to get denser, more complex circuits?

MOS model



- For $V_{gs} \leq V_T$,
  $I_{ds} = 0$
- For $V_{gs} > V_T$ and $V_{ds} \leq V_{gs} - V_T$,
  $I_{ds} = K\left[(V_{gs} - V_T)V_{ds} - \frac{1}{2}V_{ds}^2\right]$
- For $V_{gs} > V_T$ and $V_{ds} > V_{gs} - V_T$,
  $I_{ds} = \frac{K}{2}(V_{gs} - V_T)^2$

$$K \equiv \mu C_{ox}\frac{W}{L}$$

$$C_{ox} \equiv \frac{\epsilon_{ox}}{t_{ox}}$$

(Gate capacitance C_{ox} is *per unit area*)

Consequences of Scaling

All dimensions and voltages divided by the factor S(>1).

Device area	$\propto W \times L : (\downarrow S)(\downarrow S)$	$\downarrow S^2$
C_{ox}	ϵ_{ox}/t_{ox} : const$/(\downarrow S)$	$\uparrow S$
C_{total}	$\epsilon A/t : (\downarrow S^2)/(\downarrow S)$	$\downarrow S$
V_{DS}, V_{GS}, V_T	Voltages : $(\downarrow S)$	$\downarrow S$
I_d	$\mu C_{ox}(W/L)(\propto V^2)$: $(\uparrow S)(\text{const})(\downarrow S^2)$	$\downarrow S$
Slew Rate $\frac{dV}{dt}$	$I/C_{total} : (\downarrow S)/(\downarrow S)$	*const*.
Delay	$V/\frac{dV}{dt} : (\downarrow S)/(const)$	$\downarrow S$
Static Power	$V \times I : (\downarrow S)(\downarrow S)$	$\downarrow S^2$
dynamic power	$C_{total}V^2 f : (\downarrow S)(\downarrow S^2)(\uparrow S)$	$\downarrow S^2$
Power delay product	delay $\times$ power$(\downarrow S)(\downarrow S^2)$	$\downarrow S^3$
Power density	power/area : $(\downarrow S^2)/(\downarrow S^2)$	*const*.

Impact of scaling

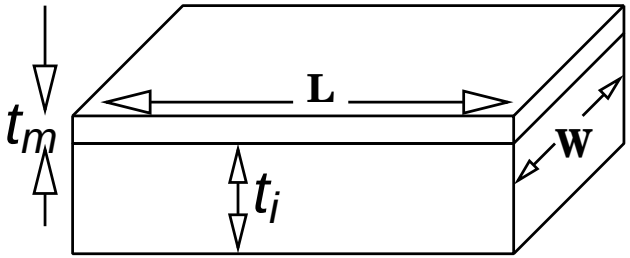
- Improved packing density: $\uparrow S^2$
- Improved speed: delay $\downarrow S$
- Improved power consumption: $\downarrow S^2$

However . . .
The above improvements apply to active circuits.

What about passive components?

Also, reduced voltages imply a lower signal to noise ratio.

Concern: Interconnect Delay

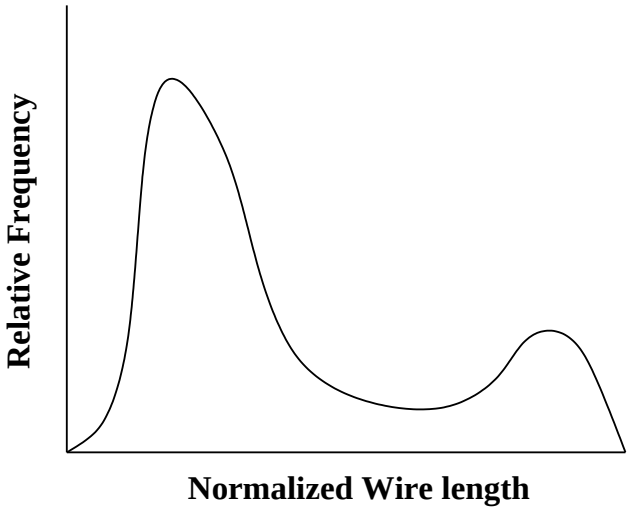


$$R = \rho \frac{L}{W t_m}, \qquad C = \epsilon \frac{LW}{t_i}$$

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

- 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, ρ, ϵ, t_m and t_i are decided by the technology).

Concern: Interconnect Delay



- Local interconnects scale with device size.
- Global interconnects scale with die size.

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

For local interconnects, L scales the same way as t_m, t_i, so delay is invariant.

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.

Buffer Insertion

Global Interconnect delay can be the determining factor for the speed of an integrated system.

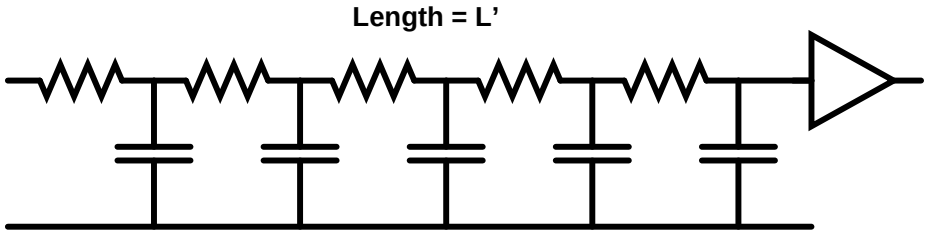
The L^2 dependence of interconnect delay is a source of particular concern.

This problem can be somewhat mitigated by buffer insertion in long wires.

We define some critical wire length and when a wire segment exceeds this length, we insert a buffer.

Repeater Insertion in Voltage Mode

What is the optimum wire length after which we should insert a buffer? (Wire Delay = $\rho\epsilon\frac{L^2}{t_m t_i} = AL^2$)



Let the wire segment length = L'.
Segment wire delay = AL'^2.
Let buffer delay = τ

For n segments, there will be n-1 buffers, and L = nL' .

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

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

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

L' should be so chosen that the wire segment delay = τ.
Total delay is proportional to n and so, is linear in L.

Difficulties with Buffer Insertion

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.
- Typically, we 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 there is active circuitry underneath, and there is no room to put in a buffer!
- We either live with buffer insertion at non-optimal wire lengths or create space by pushing out existing cells and modifying the lay out.

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!
- We need to replace buffers with bidirectional transceivers.
- These require a direction signal, which will enable a buffer in the desired direction.
- This direction signal must also be routed with the bus and should have its own buffers. It should reach the bidirectional buffers ahead of the data.

Concern: 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.

Concern: 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.

Promise of current mode signaling

- Why not signal with current rather than voltage?
- 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!

Promise of current mode signaling

If we hold the Voltage on the interconnect nearly constant

- Dynamic power is negligible.
- Latency is much lower.
- We also have the option of using multiple current levels to transmit multiple bits simultaneously. This can give
  Higher Throughput.
  Lower interconnect area.

Possibility for improving Latency, Throughput and Power **simultaneously!**

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

Digital Designers need not panic!

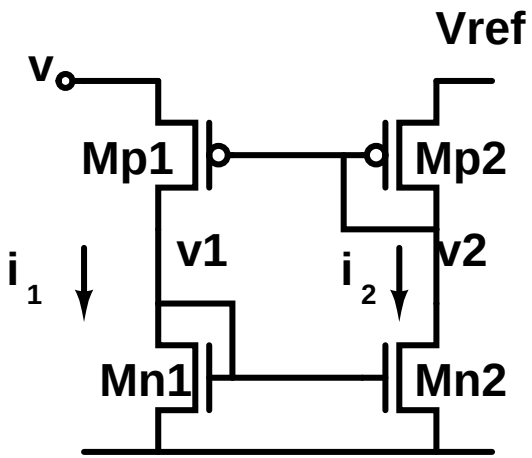
Only the interface works in current mode. Rest of the circuit is traditional.
A library circuit does the voltage mode to current conversion (transmitter) and another converts the current back to voltage mode (receiver).

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.)

Zero input impedance circuit

Low r_{in} amps are used for photo-detectors. [1]



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

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

define $\Gamma \equiv \dfrac{g_{mn2}/g_{mn1}}{g_{mp2}/g_{mp1}}$ then, $i_1(1-\Gamma) = g_{mp1}v$

This gives $r_{in} = (1-\Gamma)/g_{mp1}$

[1] C.-K. Kim et al, "High Injection Efficiency Readout Circuit for Low Resistance Infrared Detector", IEE Electronic Letters, 35, 1507, 1999.

Robustness of design

In saturation,

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

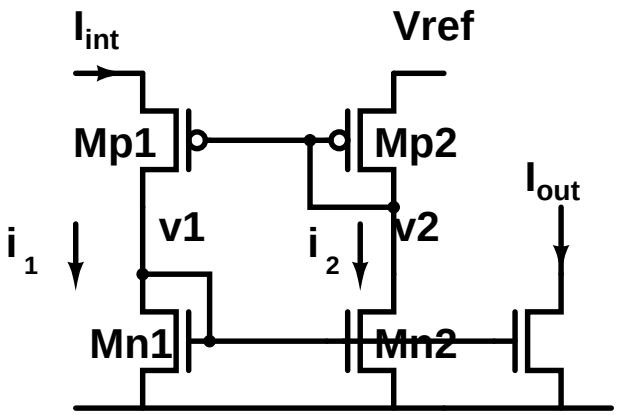
$$\text{So, } g_m = \mu C_{ox}\frac{W}{L}(V_g - V_T) = \sqrt{2\mu C_{ox}\frac{W}{L}I_d}$$

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

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

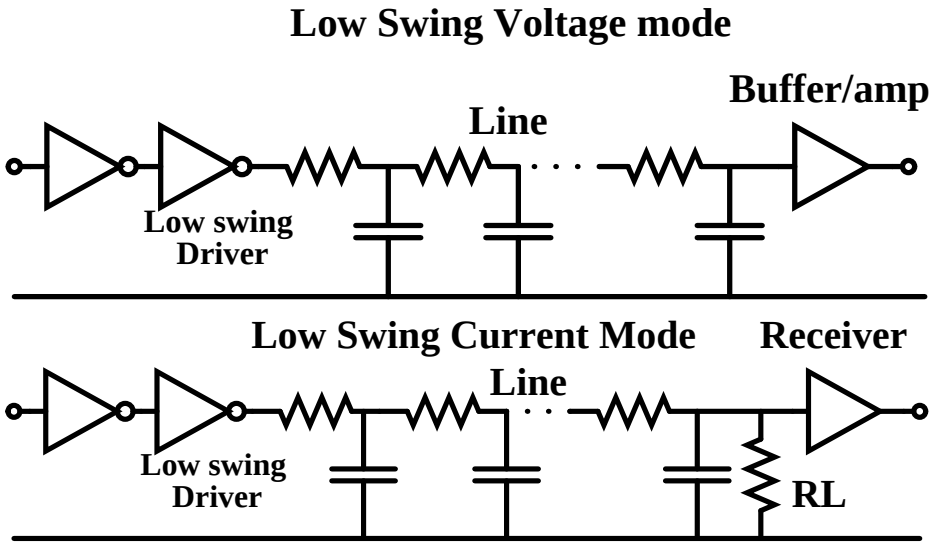
$$\text{Therefore } \Gamma \equiv \frac{g_{mn2}/g_{mn1}}{g_{mp2}/g_{mp1}} = \sqrt{\frac{(W/L)_{n2}/(W/L)_{n1}}{(W/L)_{p2}/(W/L)_{p1}}}$$

Receiver Design - Input stage



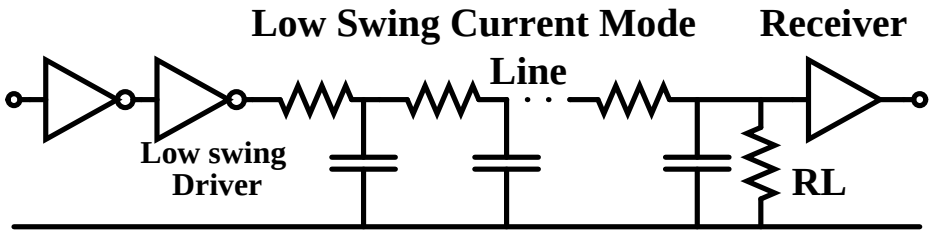
- Input resistance controlled by geometry of transistors
- Interconnect voltage held fixed
- Input resistance insensitive to process variations

Reduced swing 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.

Improving Current Mode Signaling



Current mode signaling

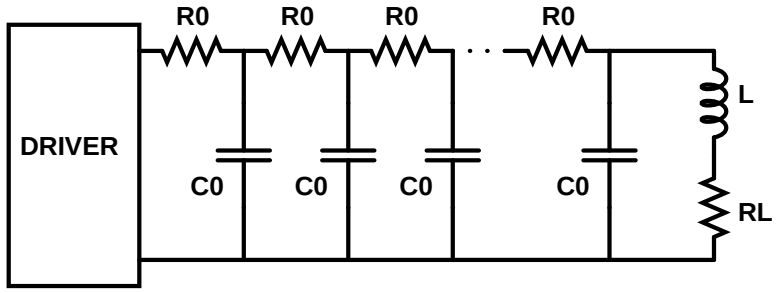
- Consumes Static Power
- Direct Trade-off between speed and static power

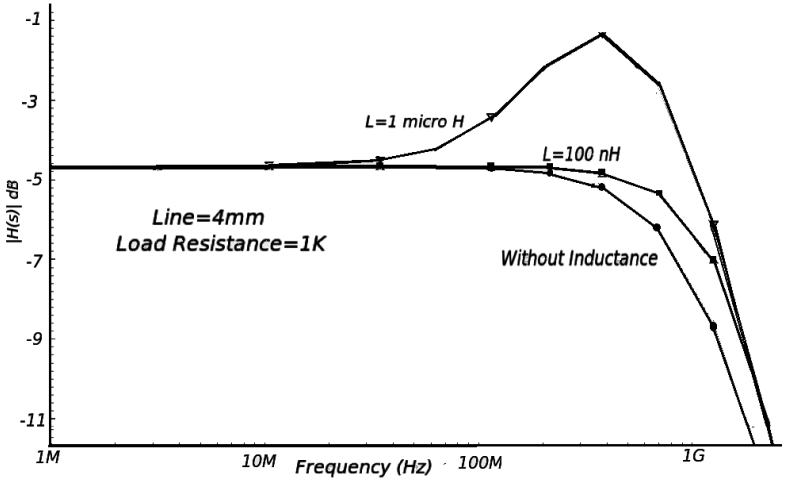
Possible Improvements

- Inductive Peaking
- Dynamic Over-driving

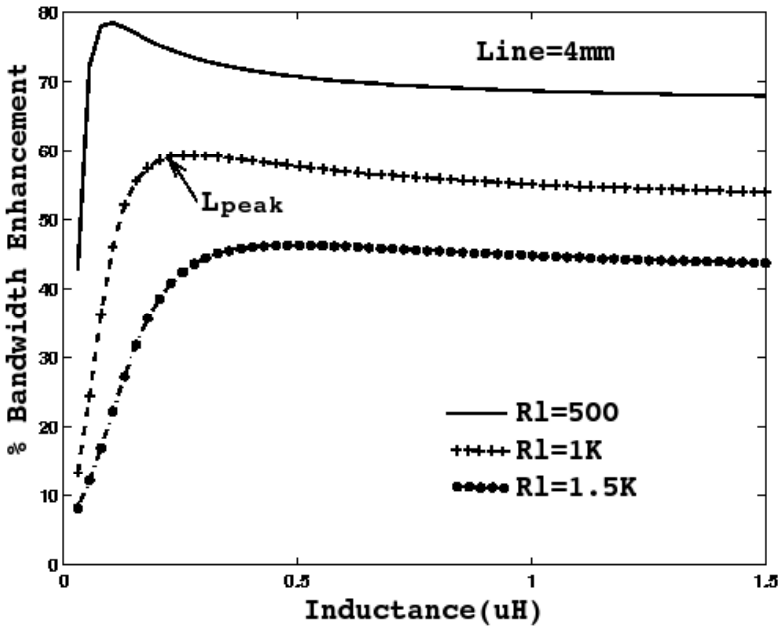
Concept of Inductive Peaking

- On-chip interconnects can be modeled as distributed RC which is essentially a low pass filter.
- Bandwidth enhancement techniques used in RF amplifiers can be employed for bandwidth enhancement on interconnects
- Inductive Peaking: Line termination circuit exhibits inductive input impedance
- Shows enhancement of about 500MHz in 3dB bandwidth.



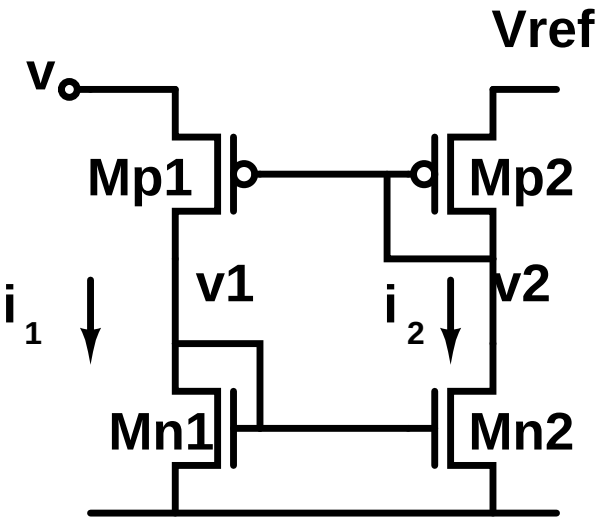


Bandwidth Enhancement Vs Load Inductance



- For a given line length, the amount of bandwidth enhancement is a function of inductance and load resistance.
- Significant bandwidth enhancement can be achieved for a wide range of inductance values greater than L_{peak}.
- The required inductance for significant enhancement in bandwidth is a few hundreds of nano Henries !!
- An active inductor is required

Beta Multiplier: A Gyrator



- 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.

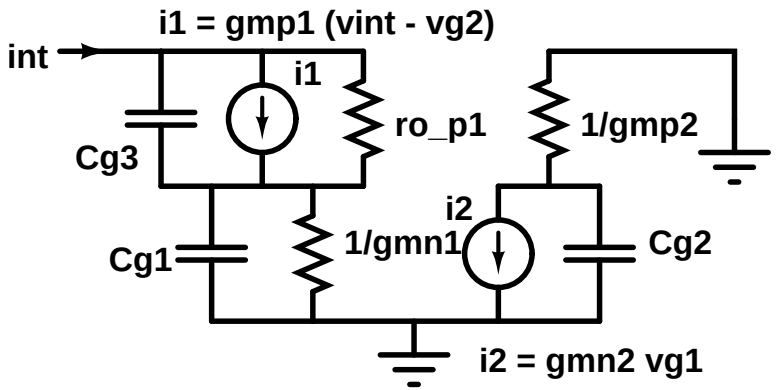
Beta Multiplier: Input Impedance

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

$$\tau_1 = \frac{C_{g1}}{g_{mn1}} \qquad \tau_2 = \frac{C_{g2}}{g_{mp2}} \qquad R_1 = \frac{1}{g_{mn1}}$$

$$\tau_3 = C_{g3} r_{op1} \qquad \tau_4 = \frac{C_{g3}}{g_{mp1}} \qquad R_3 = r_{op1}$$

$$\gamma = \frac{g_{mp1}/g_{mp2}}{g_{mn1}/g_{mn2}} \qquad k = \frac{R_1}{R_3}$$



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

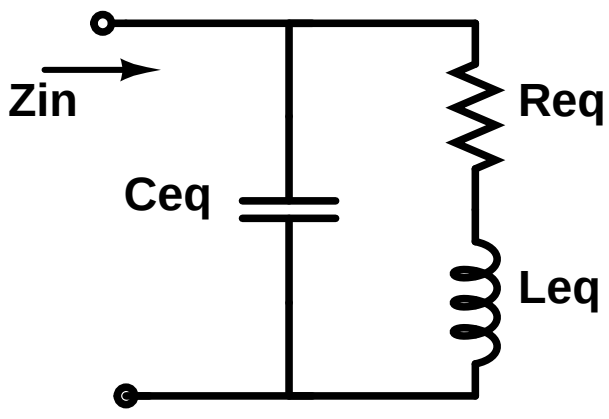
Beta Multiplier : Equivalent Circuit

- Relative location of poles and zeros determine nature of impedance (inductive of capacitive)
- If the first zero occurs a decade prior to the first pole, input impedance is inductive
- $\gamma - \frac{1}{g_{mn1} r_{op1}} > 0.9$ and any two time constants being equal ensures that a zero occurs a decade prior to the first pole

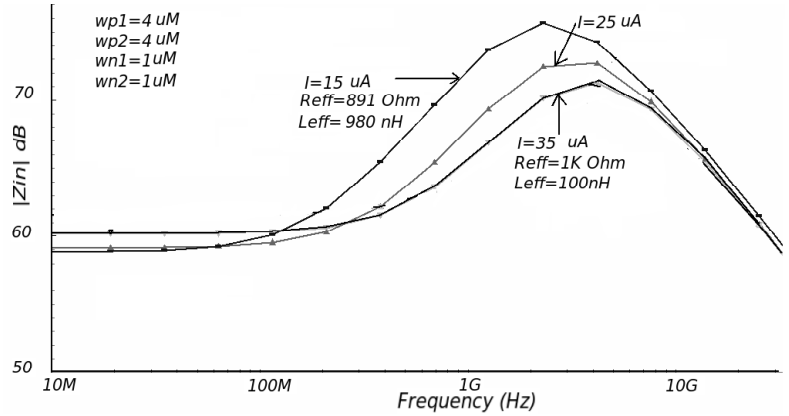
$$L_{eff} = \frac{r_{op1}}{g_{mp1} r_{op1} + 1} \left\{ \frac{C_{g1}}{g_{mn1}} + \frac{C_{g2}}{g_{mp2}} + \frac{C_{g2}}{g_{mp2} g_{mn1} r_{op1}} + \frac{C_{g3}}{g_{mn1} g_{mp1} r_{op1}} \right\}$$

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

$$C_{eff} = K C_{gx}$$

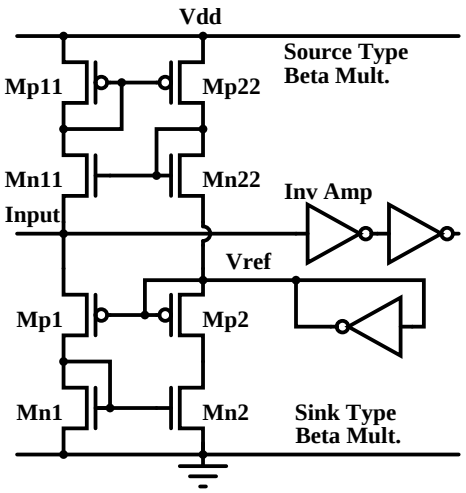


Beta Multiplier : Input Impedance Control



- Beta Multiplier shows an effective inductance 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



- Effective impedance offered by the receiver is equal to the parallel combination of the impedance offered by individual beta multipliers.
- Voltage at input node swings around V_{ref}. 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 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.

Simulation Results

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

Signaling Scheme	Delay (ps)	Throughput (Gbps)	Power (μW)	Area (μm^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

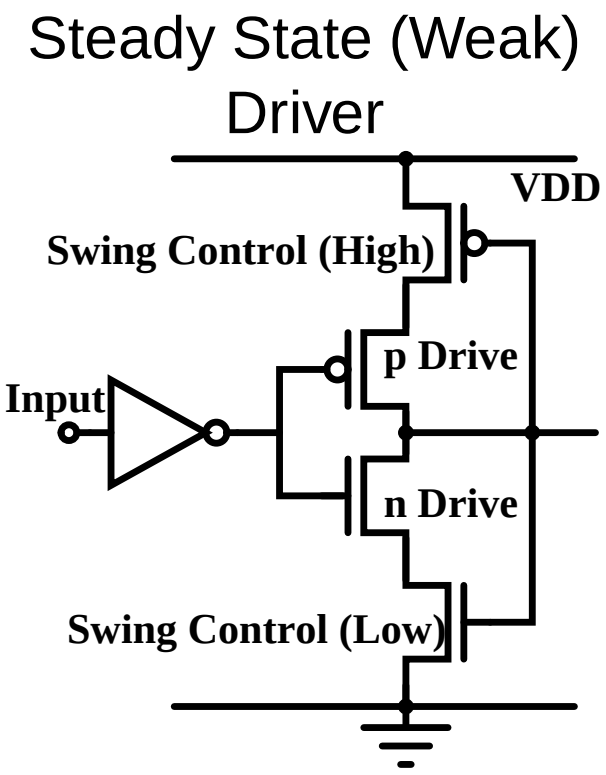
- Inductive termination gives 16% improvement in delay and about 18 % improvement in power. Also more than 50 % improvement in delay at the same time an order of magnitude lower power.

[1] M Dave et. al., ISLPED 2008, [2] V. Venkatraman et. al. ISQED 2005

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 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

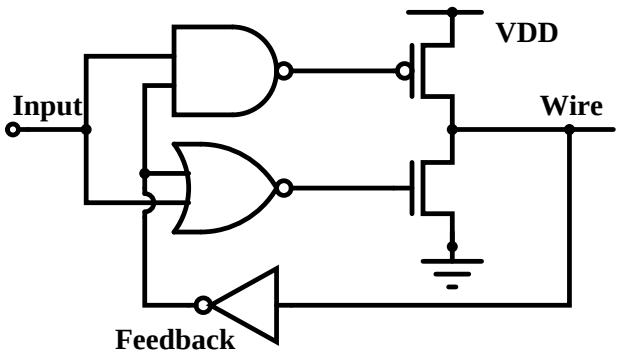


A. Katoch et. al. ESSCIRC, 2005

- The p channel driver gate is low (enabled) when the input is 1.
- As the line reaches $V_{DD} - V_{Tp}$, the upper p channel transistor turns off, restricting line voltage swing.
- Similarly the n channel driver transistor is enabled when the input is 0 and the lower transistor turns off when the input approaches V_{Tn} during discharge.

Possible implementation of Dynamic Overdriving

Dynamic (Strong) Driver

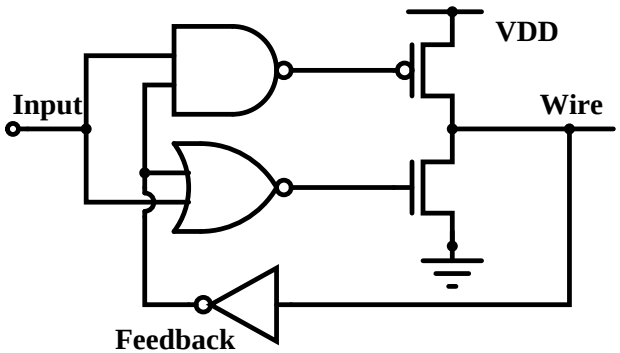


The feedback inverter acts as an inverting amplifier converting low swing logic levels on the wire to full swing (inverted) CMOS logic level on its output.

- P channel gate is low (enabled) only when the input is high AND the line is at 0.
- N channel gate is high (enabled) only when the input is low AND the line is at 1.
- Input to the feedback inverter is a low swing level around $V_{DD}/2$. Therefore it consumes static power.

Self limiting Strong Driver

Dynamic (Strong) Driver



- Input = 1, Wire voltage $< V_m$

  Inverter output = 1, NAND output = 0, NOR output = 0

  P channel driver dumps current to charge the line.
- Input = 0, Wire voltage $> V_m$

  Inverter output = 0, NAND output = 1, NOR output = 1

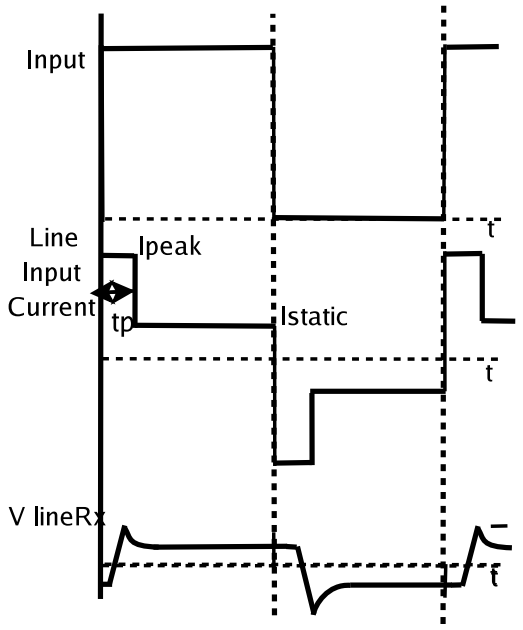
  N channel sinks current to discharge the line.

- As soon as low swing logic level on the line = input
  Inverter output = $\overline{input}$, NAND output = 1, NOR output = 0
- This disables both drive transistors automatically.

A. Katoch et. al. ESSCIRC, 2005

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?

Current Mode Signaling Schemes with Ideal Components

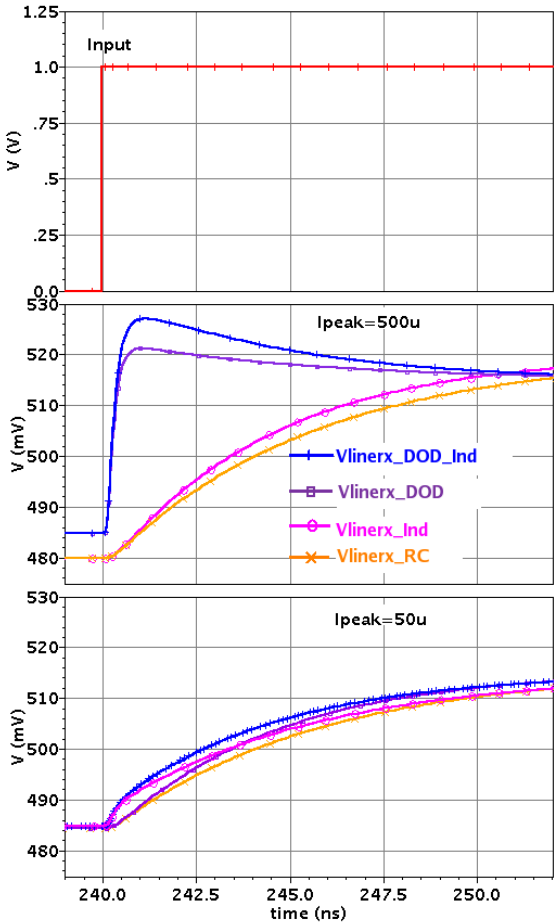
Following four current mode signaling schemes were simulated:

- 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

Implementation details of these circuits are:

- Dynamic Overdriving driver is implemented by ideal VCCS with current wave shape as shown in the figure. Controlling voltage is input.
- Simple driver is implemented as VCCS with square wave shape. The input current ranging from $-I_{avg}$ to $+I_{avg}$.
- $I_{avg} = \frac{I_{peak} t_p + I_{static}(t - t_p)}{t}$
- $R_L = 4k\Omega, l = 4\mu H$

Comparison of Delay



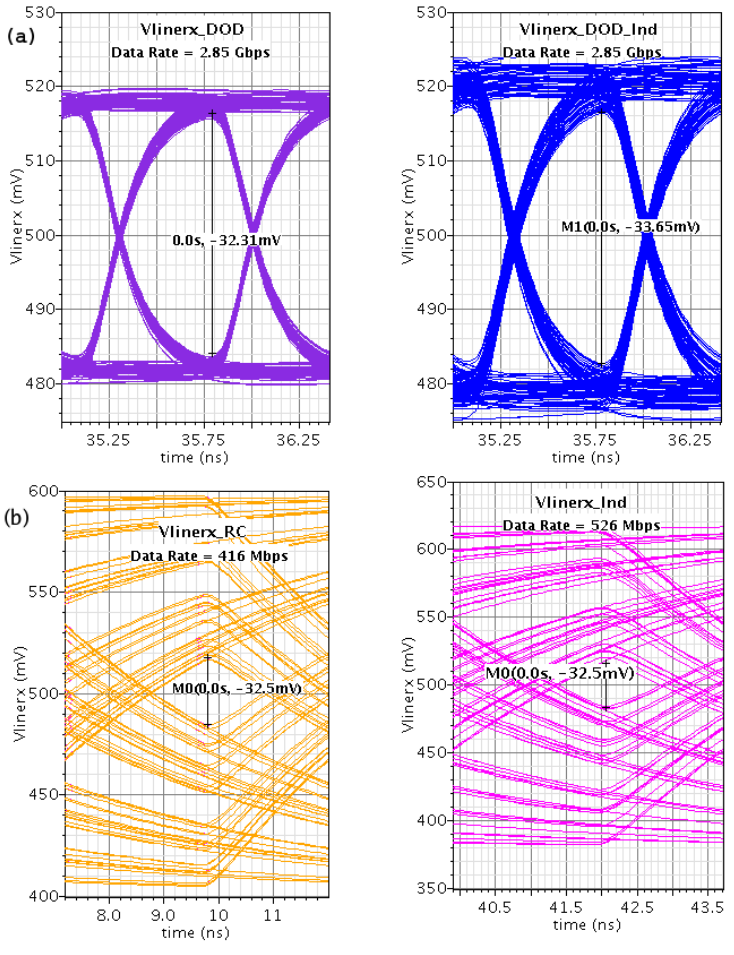
With Large Overdrive ($I_{peak} = 500\mu A$)

- 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)



- 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 over RC
- For low power and low data rate applications, the use of inductive peaking can give 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 throughput by 20%

Part II

Variation Tolerant Current Mode Signaling

Need for Process Variation Tolerance

Effect of Process Variations on different CMS Schemes

The Proposed Variation Tolerant CMS Scheme

Performance Evaluation

Bidirectional Links

- Simulated Performance of Bidirectional Link

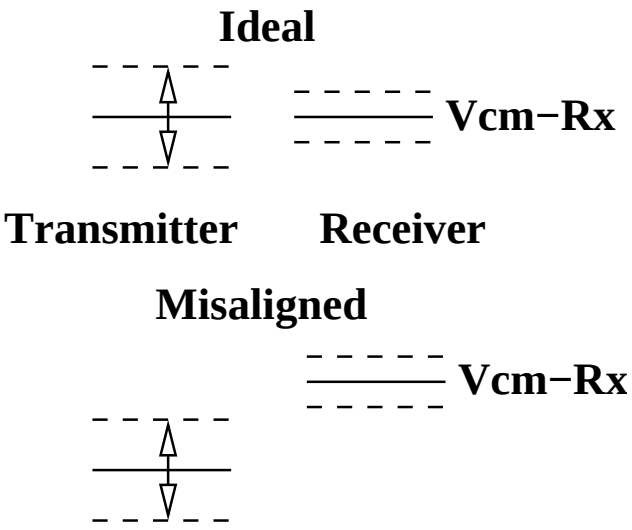
Need for Process Variation Tolerance

- 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 60%.
- 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.
- Intra-die variations cause mismatch in parameters of transmitter and receiver transistors.

Robustness requirements

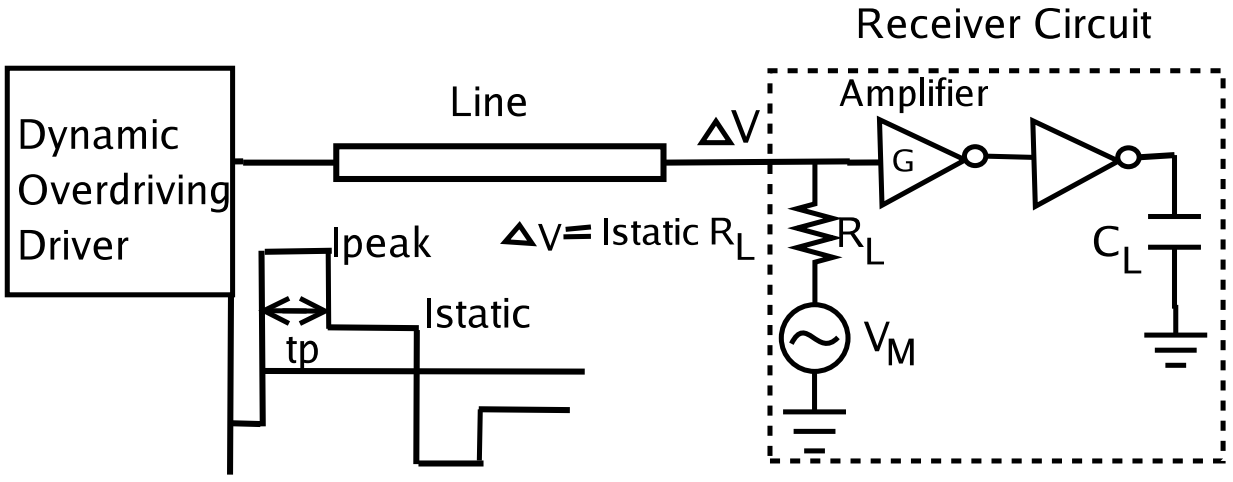
- Process, Supply Voltage and Temperature variations will affect the core logic as well as data communication circuitry.
- The requirement for data transmission is therefore not of *complete invariance* with respect to PVT variations.
- We have to ensure that throughput and delay properties of the interconnect 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.
- Because global interconnects, by definition, connect remote points on the die, on chip variations can be of greater concern.

Effect of common mode voltage mismatch



- In case of ideal match, small 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.
- It is better to have smart bias circuits, which will reduce mismatch and the need for a large swing.

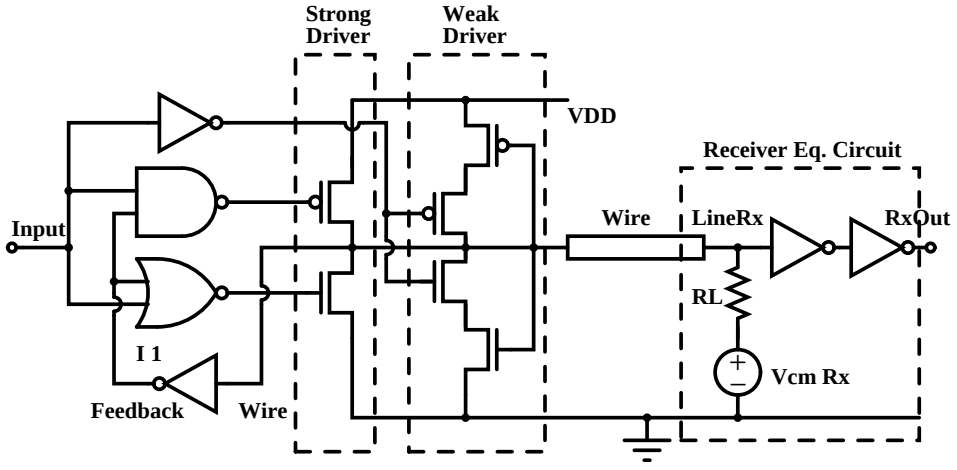
System parameters affected by variations



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. ΔV: Line voltage swing at the receiver end in steady state
4. Mismatch between any V_{CMRx} and operating point of an amplifier

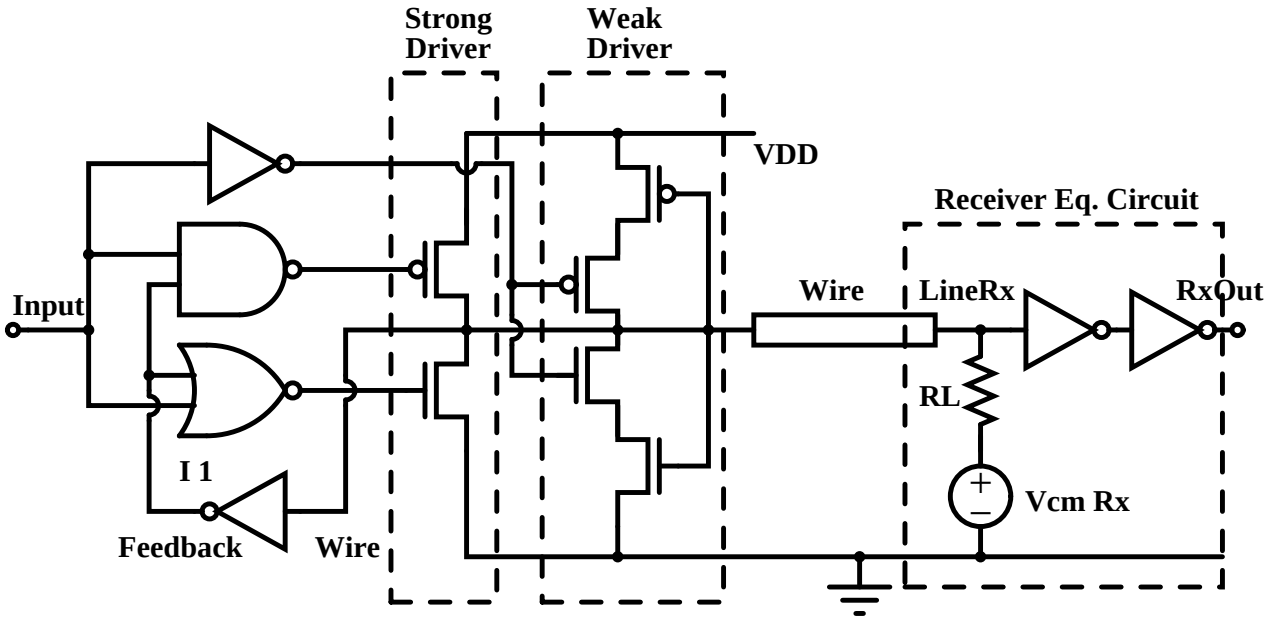
CMS Scheme with Feedback (CMS-Fb)



- NAND/NOR generates pulses to turn-on/off the strong driver
- Input transition $\rightarrow$ the strong driver turns on
  $\rightarrow$ line voltage at transmitter end crosses V_M of inverter I1
  $\rightarrow$ strong driver turns off.
- Weak driver supplies I_{static} and line voltage swing at receiver end is $V_{CMRx} \pm I_{static} R_L$

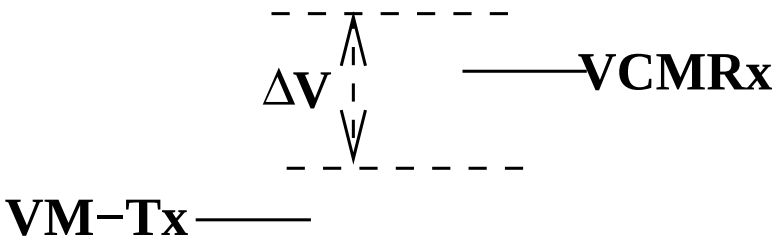
A. Katoch et. al. ESSCIRC, 2005

Effect of Inter-die Process Variations on CMS with feedback

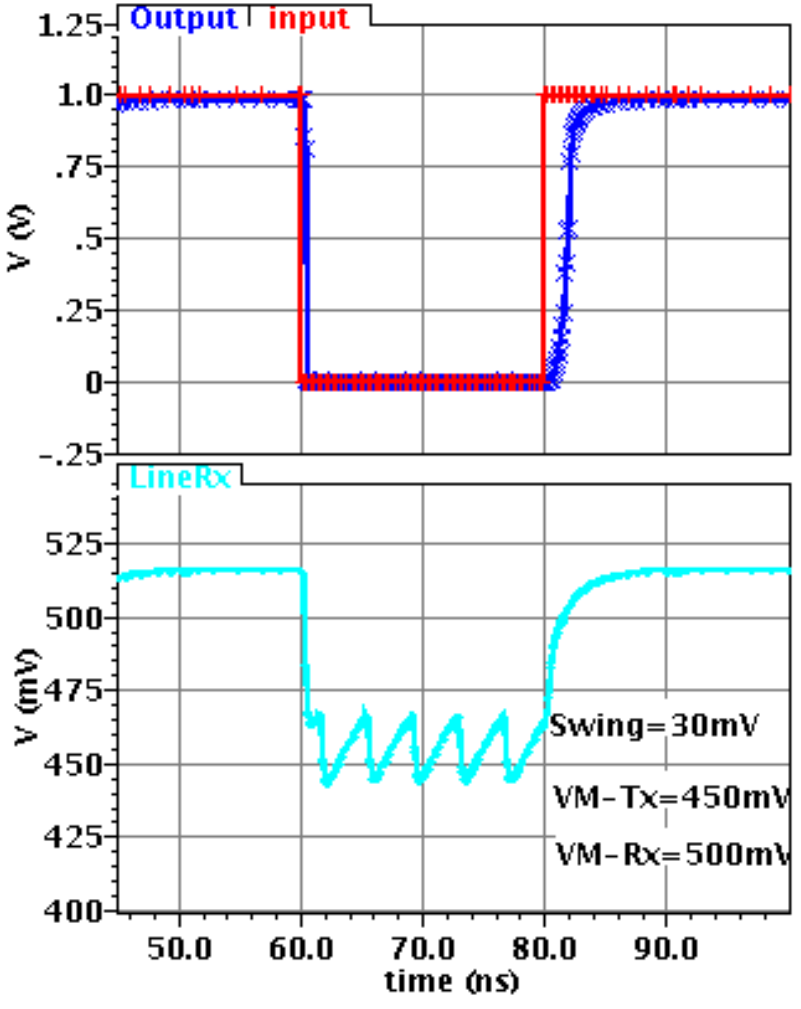


- Variations in I_{peak} are well compensated due to the feedback at the driver end.
- If the driver is weaker due to process variations, the feed back system keeps it on for longer till the line reaches the desired voltage.
- This might, however, not be optimum from a power point of view.

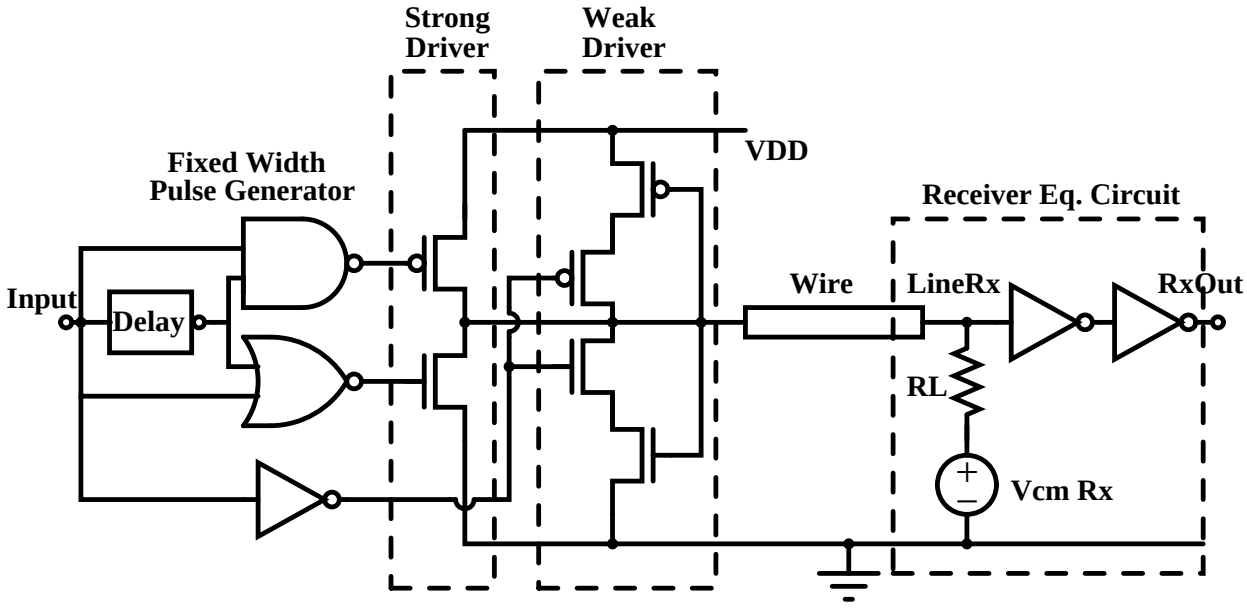
Effect of Intra-die Process Variations on CMS-Fb



- Line voltage is not constant for constant low input voltage
- During low to high transition the strong driver is turned off well before the line voltage crosses V_{CMRx}



CMS Scheme without Feedback (CMS-Fpw)

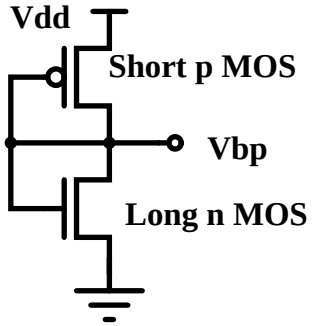


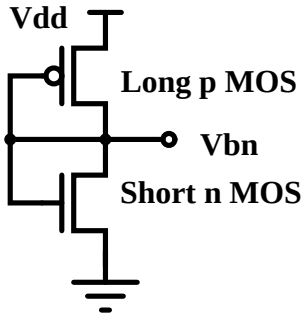
- t_p is given by delay element
- Less sensitive to intra-die variations
- In the skewed corners, sourcing I_{peak} and sinking I_{peak} are different, leading to different rise and fall delay
- Throughput can degrade significantly in skewed corners

A.Tabrizi et. al. MWSCAS, 2007

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.

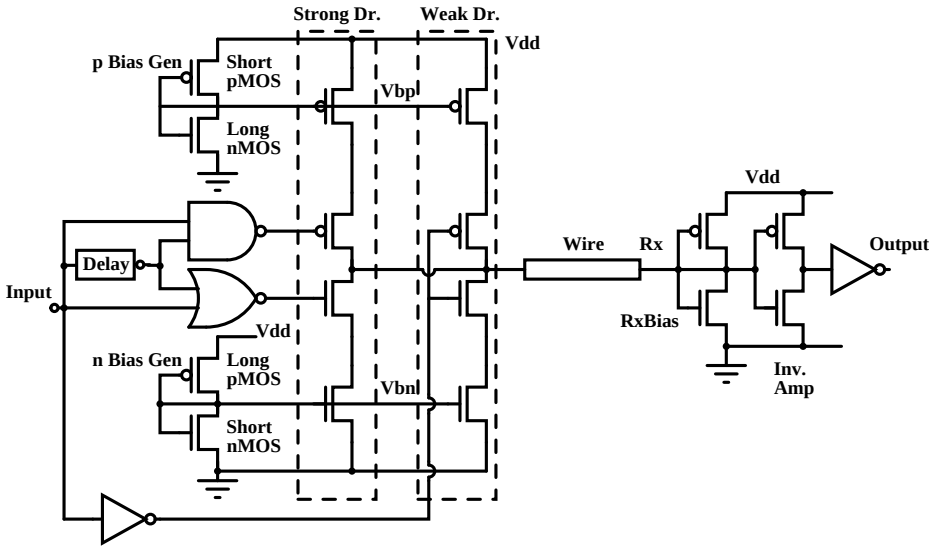




- Long Channel transistors show relatively less variation with process compared to Short Channel transistors in the same process.
- We can make use of this difference to design a bias generator which senses the process corner and tries to increase the transistor current in the slow corners and to decrease 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.

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 VM of inverter can be up to 40 mV. [2]. For VM-mismatch of 40 mV

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

[2] 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

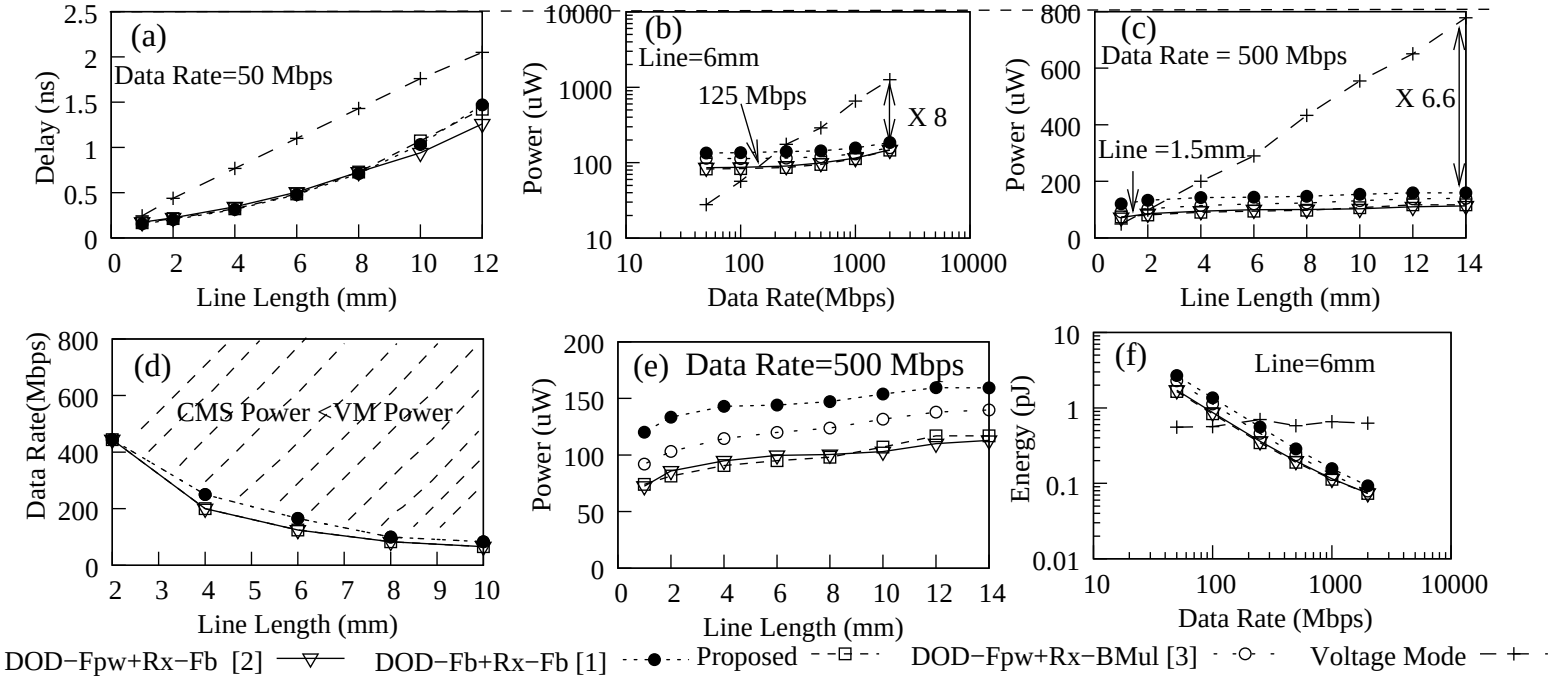
Overall Comparison

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

Signaling Scheme	Delay (ps)	Throughput (Gbps)	Power (μW)	Area (μm^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

Overall Comparison



Bidirectional Links

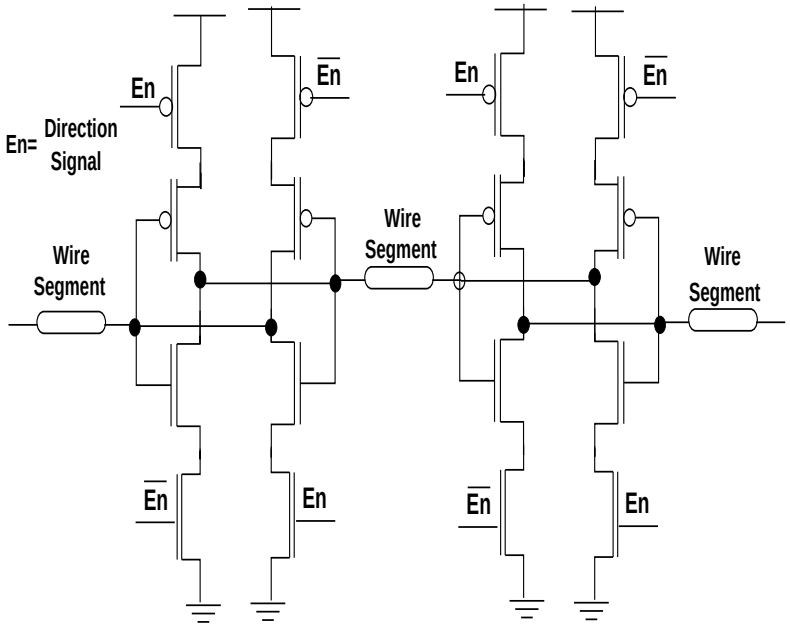
In many applications, on-chip buses need to carry signal in both directions.

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



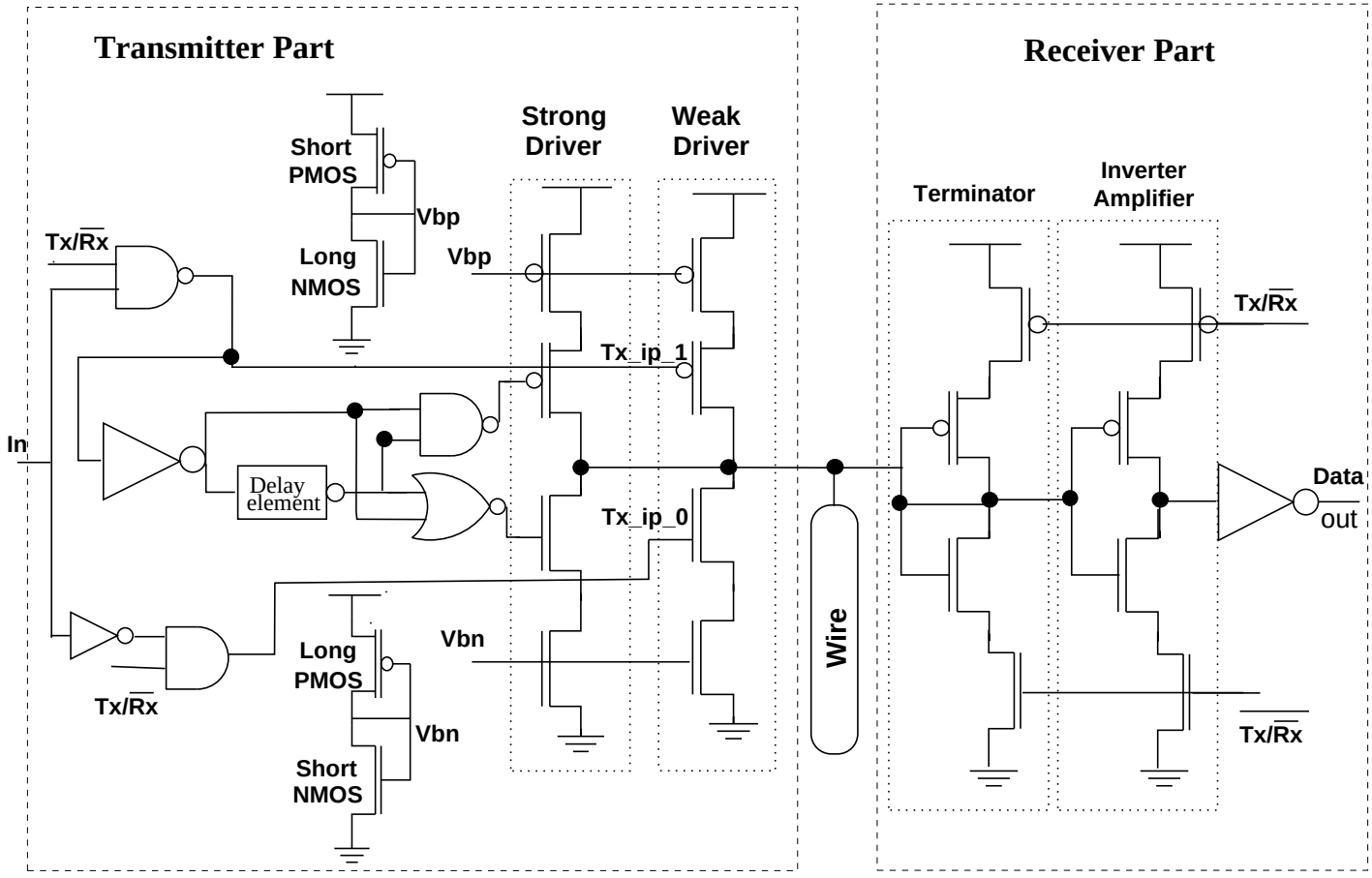
- One of the two tristate buffers is enabled at a given time
- Two transistors in stack $\Rightarrow$ increased sizes of PMOS and NMOS
- Delay of a bidirectional repeater is more than that of a unidirectional buffer
- Direction control signal is required by each repeater
- Buffers offer huge load to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

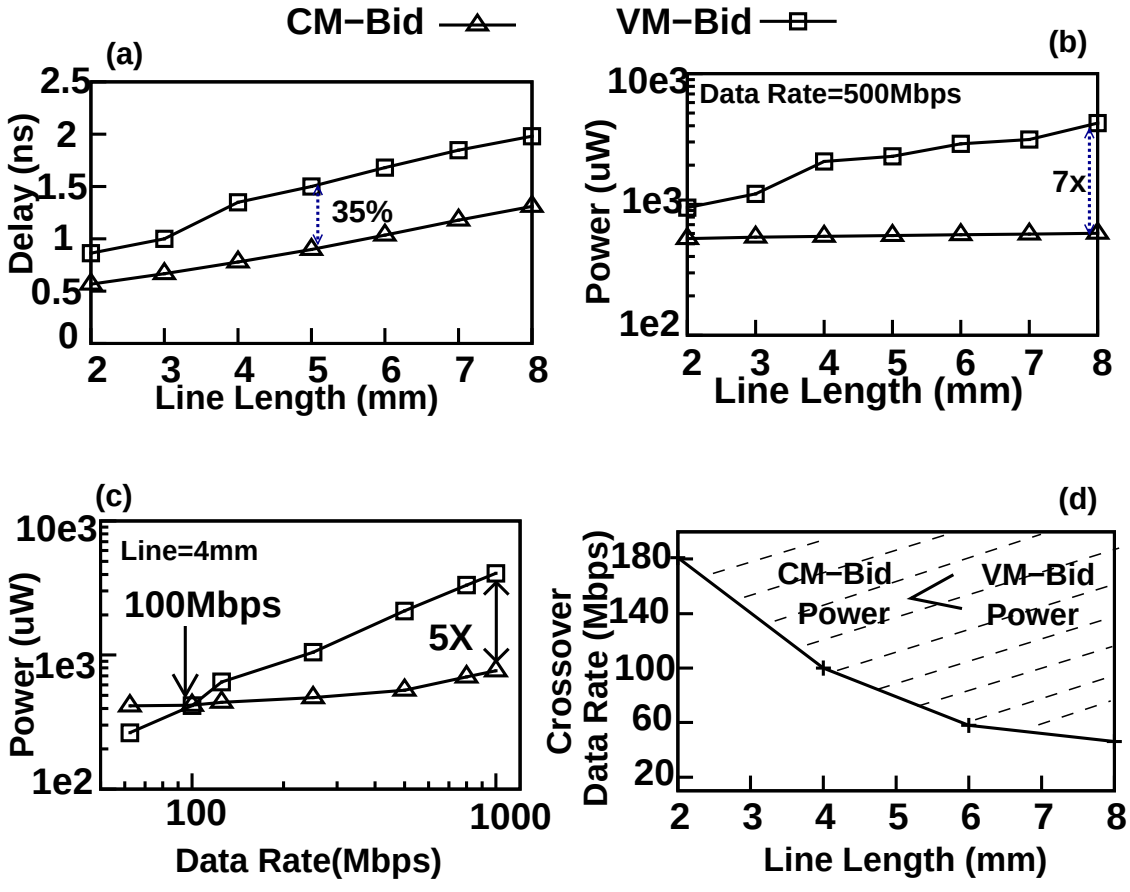
Proposed Current-Mode Transceiver



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

Speed-Power of Proposed Bidirectional CMS Scheme

Current-Mode Vs. Voltage-Mode



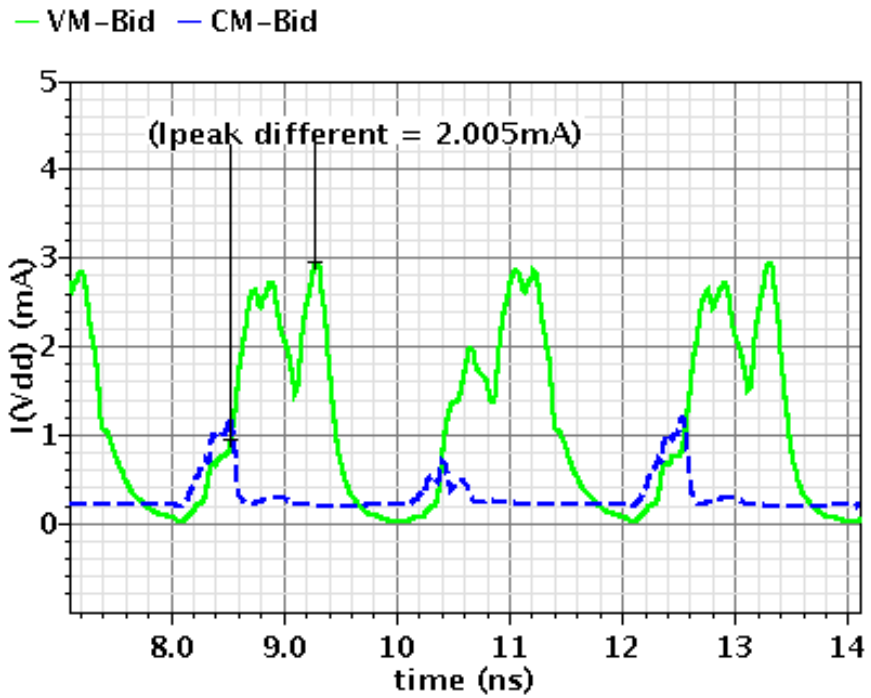
- 35% improvement in delay for nearly all line lengths
- 1.7× lower power for 2mm lines and 7× lower power for 8mm line
- Power crossover frequency 100Mbps for 4mm long lines
- 5 × reduction in power at 1Gbps
- For lines longer than 2mm communicating at data-rates more than 180Mbps, the proposed scheme consumes less power than voltage-mode

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

Line Characteristics: R=211Ω/*mm* and C=0.245pF/mm

Effect on Supply Noise

Peak Current Drawn From Supply



- 68% reduction in peak current and hence contribution to supply noise is much less
- 80% reduction in active area

Performance of Proposed Scheme in Four Digital Process Corners

Specs	Delay (ns)		Power (μW)	
	VM-Bid	CM-Bid	VM-Bid	CM-Bid
TT	1.35	0.81	2127	567
SS	1.57	0.90	2055	435
FF	1.21	0.69	2163	727
FNSP	1.35	0.80	2113	572

For a 4mm line operating at 500Mbps

- 38% improvement in delay even in worst case (SNFP)
- 3.45$\times$ lower power consumption even in worst case (SNSP)

Part III

Implementation on Si and Measured Results

On chip measurement
- Time to Frequency Conversion
- Time to Voltage Conversion

Implementation on a Test Chip

Measurement Results
- Bidirectional Lines

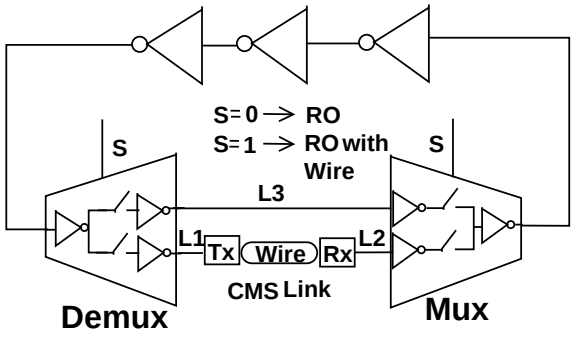
Motivation

- Delays of on-chip interconnects are of the order of hundreds of pico-seconds.
- It is nearly impossible to measure these off-chip.

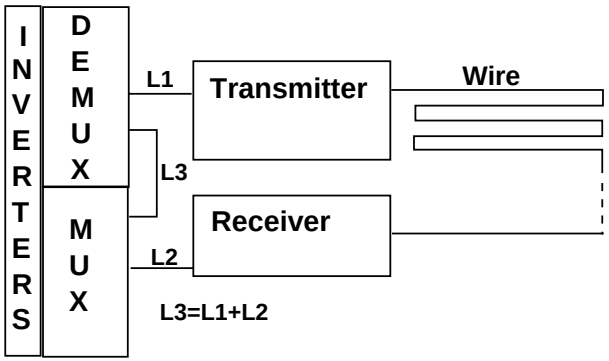
We need on chip delay measurement circuits. We have designed two test circuits based on:

- Time to Frequency Conversion
- Time to Voltage Conversion

Time to Frequency Conversion

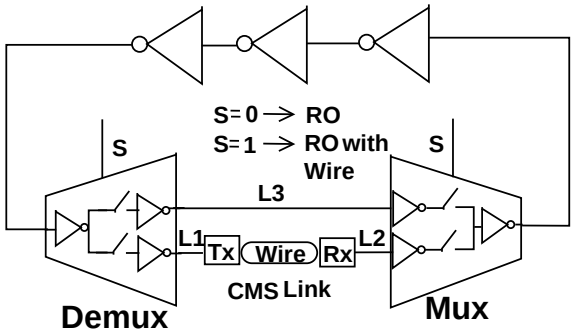


(a) Delay Measurement Circuit: Principle

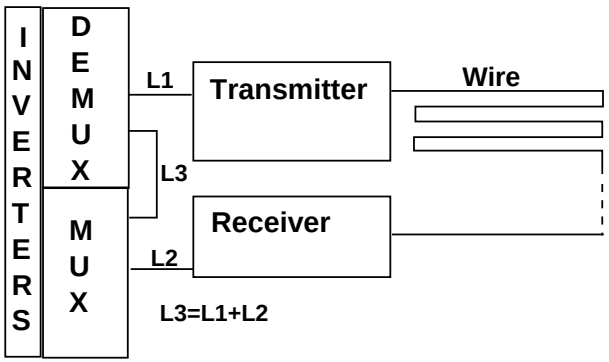


(b) Delay Measurement with CMS Link: Floorplan

- Transmission gates were used to implement switches.
- Multiplexer(demultiplexer) are designed so that delays for both possible paths through the mux/demux pair are the same.
- The floor plan of the circuit is such that the beginning and the end of the long interconnect are close to each other.
- Therefore when the short path L3 is chosen, the total delay corresponds to the delay in inverters, mux/demux etc.



(a) Delay Measurement Circuit: Principle



(b) Delay Measurement with CMS Link: Floorplan

- We first measure the frequency of oscillation choosing the short wire path between the demux and mux.
- This gives the delay of the measurement circuit except for the system under test.
- We now select the interconnect system whose delay we want to measure and find the frequency again.

$$Delay = 0.5 \left\{ \frac{1}{f_{RO}} - \frac{1}{f_{system}} \right\}$$

Time to Frequency Conversion: Accuracy

To assess the accuracy of the scheme, we simulated the whole circuit, for different line lengths up to 14 mm in a 180 nm process.

- The delay through the interconnect scheme was noted from the simulation results. We call this the “Simulated Delay”
- The delay was also calculated by the formula:

$$0.5 \left\{ \frac{1}{f_{RO}} - \frac{1}{f_{system}} \right\}$$

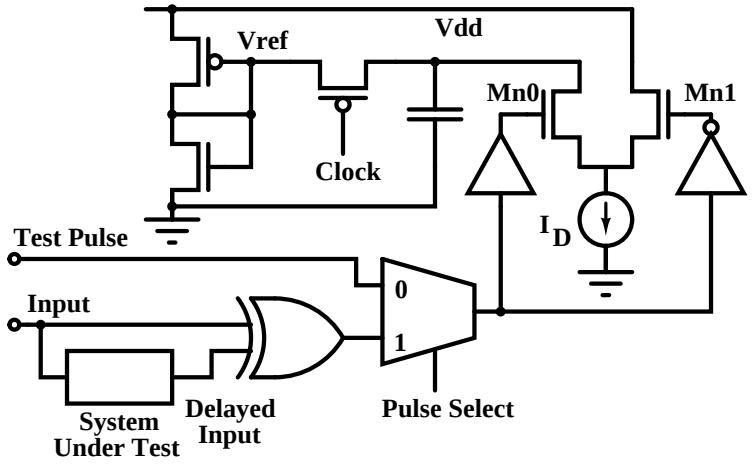
  We call this the “Calculated Delay”
- These results were tabulated to assess the expected accuracy from this test scheme.

Time to Frequency Conversion: Accuracy

Line Length (mm)	Simulated Delay (ps)	Calculated Delay (ps)	% Error
4	501	507	1.2
6	661	658	0.4
10	1068	1077	0.8
14	1575	1599	1.5

- Delays are the average of rise and fall delay
- Power-delay product can be evaluated using this circuit.
- This being a differential measurement, the only source of error is differences in rise and fall time

Time to Voltage Conversion



- Capacitor C is pre-charged to peak value during the negative phase of the clock.
- It is then discharged for a time equal to the delay through the system.
- $Delay = \frac{C\Delta V}{I} = k\Delta V$
- Value of k is found experimentally using a calibration pulse of known duration.

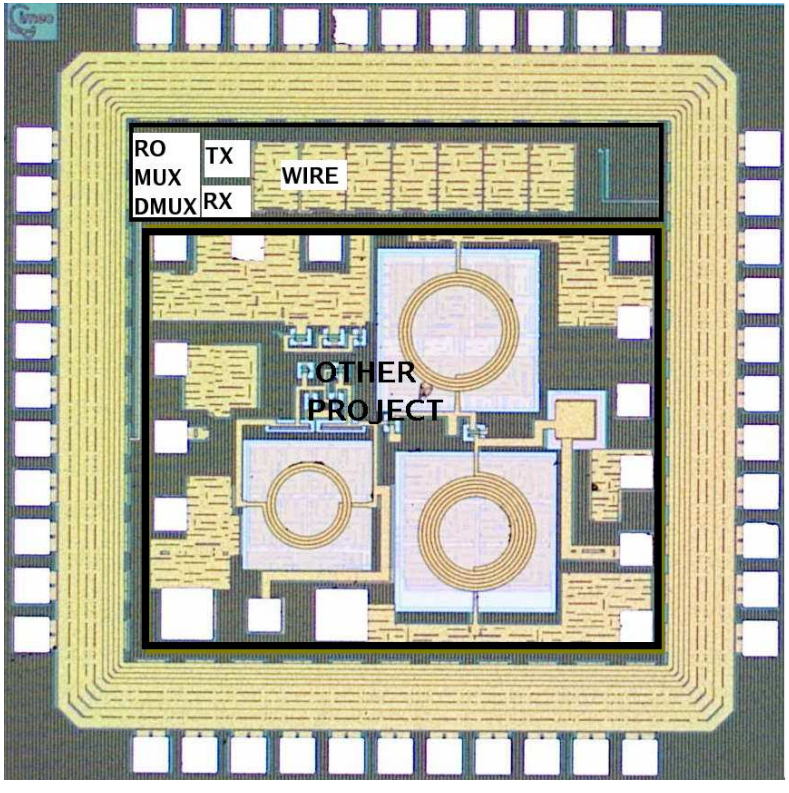
Time to Voltage Conversion: Accuracy

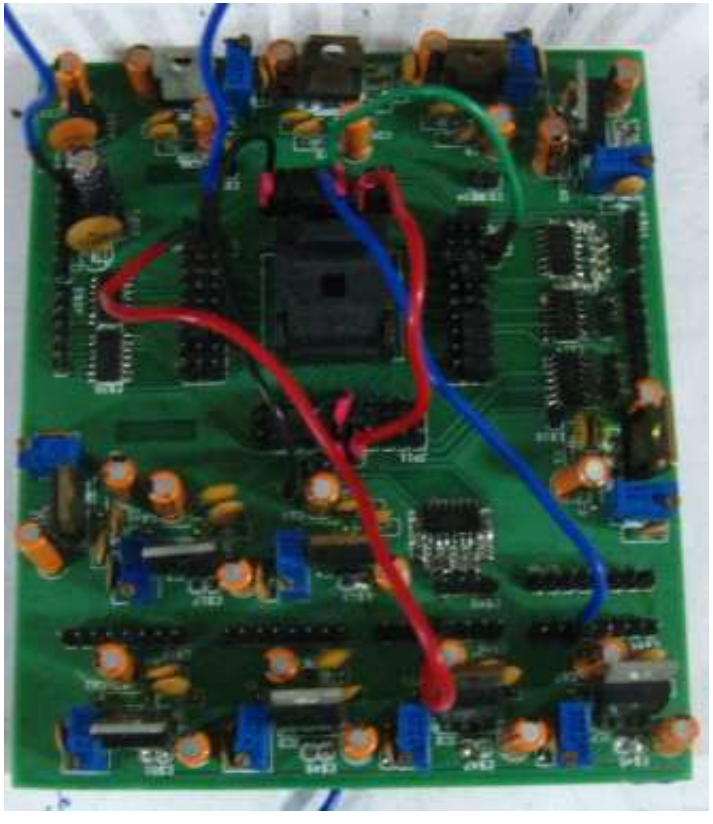
Line Length	Simulated Delay		Calculated Delay		Error	
	rising	falling	rising	falling	rising	falling
(mm)	(ps)	(ps)	(ps)	(ps)	%	%
4	380	393	378	398	0.8	1.0
6	478	497	482	503	0.8	1.2
10	730	769	733	781	0.4	1.8
14	1065	1149	1078	1171	1.2	1.9

- This scheme permits the measurement of rise and fall delays separately.
- Accuracy of about 2% is predicted by simulations.

Current-Mode Signaling Test Chip

- 1.5mm $\times$ 1.5mm chip fabricated in 180nm MM/RF process
- 44-pin die packaged in QFN56 package





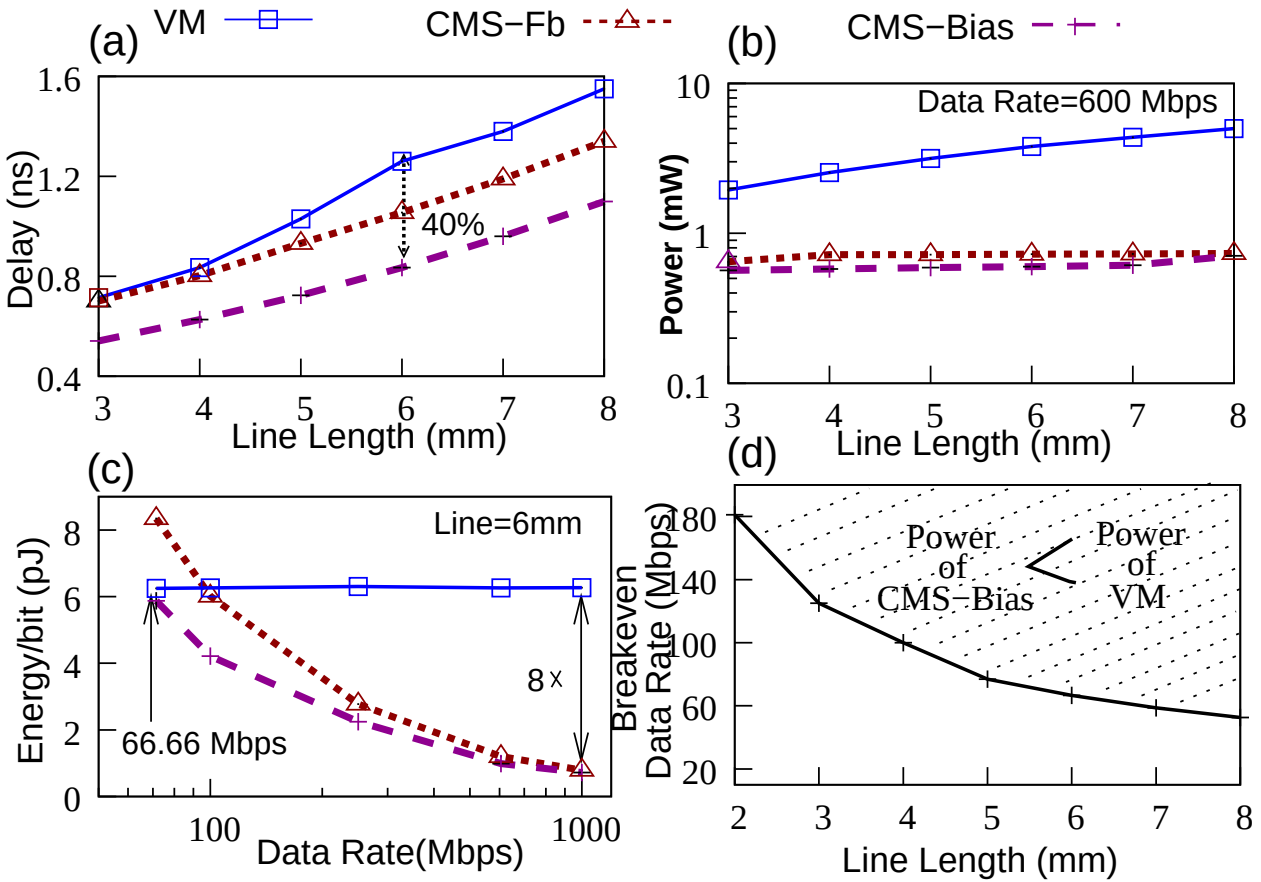
Measurement Results

(Frequency measured using a 6-digit frequency counter)

Signaling Scheme	Delay (ns)	Energy (pJ)	EDP (pJ$\times$ns)	Measured at Data Rate (Mbps)
Voltage Mode	1.191	4.54	5.328	371
CMS-Fb	1.006	1.52	1.52	400
CMS-Bias	0.938	0.851	0.799	621

The proposed circuit offers 22% improvement in delay and 85% improvement in EDP over voltage-mode scheme.

Performance of Proposed CMS Scheme



- At least 7× lower power in the worst process corner
- 78% gain in active area
- 65% reduction in peak current

Voltage-mode scheme was optimized for delay separately for every line length

Comparison with Existing Dynamic Overdriving CMS Schemes

Source	JSSCC 2006	CICC 2006	ESSCIRC 2005(CMS-Fb)	This work	This* work
Sim./Measured	Meas.	Meas.	Meas.	Meas.	Sim.
Tech.	130nm	250nm	130nm	180nm	180nm
Line (mm)	10	5	10	6	6
Gain in Delay	32%	28.3%	53%	22.5%	32%
Gain in Energy/bit	35.48%	67%	25%	**81.0%**	87%
Gain in EDP	56.5%	76.8%	65.5%	**85%**	90%
Data Rate (Gbps)	3	2	0.7	**0.62**	1
Activity α	1.0	1.0	NA	1.0	1.0

Comparison With Voltage Mode Buffer Insertion

- The proposed dynamic overdriving CMS scheme offers 26-40% improvement in delay over the voltage-mode scheme for 2mm-8mm long lines.
- These also offer improvement in energy consumption over buffer insertion scheme for lines longer than 2mm operating at data-rates more than around 66Mbps.
- The proposed 6mm long link reduces energy consumption at least by a factor of 7 compared to the voltage-mode scheme at 1Gbps.
- It offers 85% improvement in Energy Delay Product (EDP) over voltage-mode scheme.

Comparison With Other Current Mode Schemes

- The scheme proposed by us offers 22% improvement in Power Delay Product (PDP) over the current mode scheme with feedback proposed by Katoch et al.
- The CMS scheme with feedback is sensitive to intra-die variations. Our CMS scheme remains faster than logic circuit even in the presence of intra-die and inter-die process variations.

Measurement Results for Bidirectional Links

- Measurement results match simulation results within 20%
- Voltage-mode bidirectional link was not put on silicon due to limited number of pads

Signaling Scheme	Delay (ns)	Power (μW)	PDP (mW$\times$ns)	Data rate of Measurement(Gbps)
CM-Bid	1.16	680	0.788	0.56

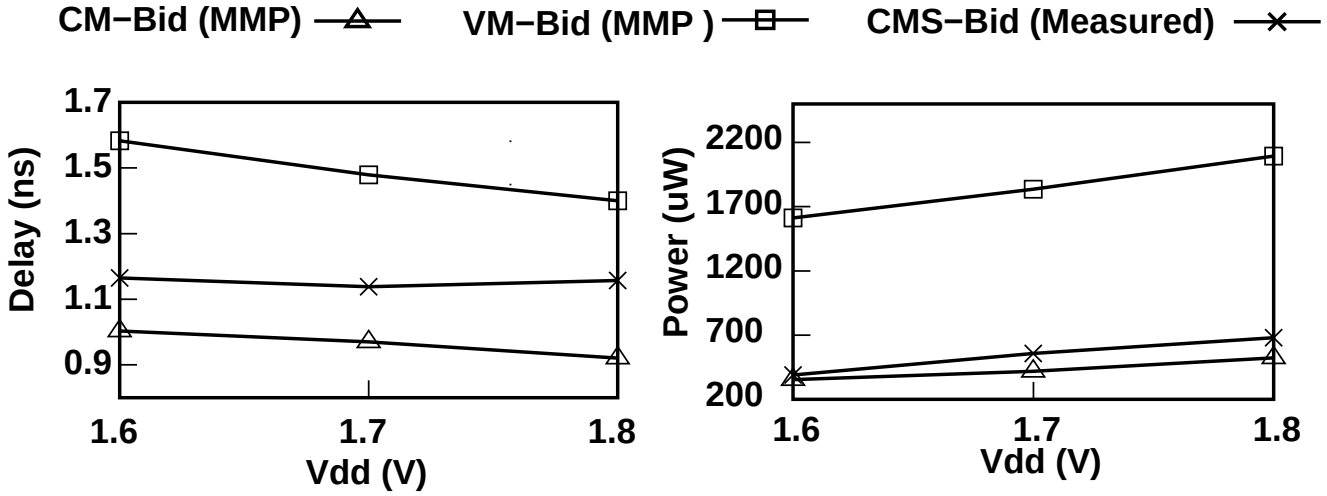
Matched Model Parameters

- BSIM parameters corresponding to this run were extracted
- A few main model parameters (BSIM) were changed to define four process corners (FF,SS,FS,SF)
- Main model parameters (BSIM) were adjusted to match I_{sat}, V_{th}, I_{off} and a few points on measured I_{ds}-V_{gs} characteristics of the devices fabricated in this process run.

Simulation with Matched Model Parameters

Parameters	TT	Measured	MMP	% Match
	Basic Device Parameters			
I_{satn}(mA)	6.23	6.44	6.43	99.8
I_{satp}(mA)	2.40	2.22	2.28	97.3
V_{tn}(mV)	501	510	506	99.2
V_{tp}(mV)	494	493	499	98.8
I_{offn}(pA)	75	170	120	82.4
I_{offp}(pA)	80	48	58	80.5
$\mathbf{I}_{dsn}/\mathbf{I}_{dsp}$@ V_{gs}	$\mathbf{I}_{ds} - \mathbf{V}_{gs}$ **points**			
I_{dsn}@0.9 (μA)	66.6	65	66.4	97.85
I_{dsp}@0.9 (μA)	76.2	70	67.5	96.45
I_{dsn}@1.2 (μA)	154.4	150	145	96.67
I_{dsp}@1.2 (μA)	191	170	172	98.82
I_{dsn}@1.8 (μA)	347	330	317	96
I_{dsn}@1.8 (μA)	491	440	452	97.27

Measurement Results and Simulation Results with MMP



Improvement in Specs For Simulations using MMP

Vdd (V)	Delay(%)	Power(x)	PDP(x)
1.6	36.8	4.5	7.2
1.7	34.4	4.39	6.8
1.8	34.21	4.01	6.0

Conclusion

- Global interconnects form a major bottleneck for performance of digital system at scaled down technology.
- Use of current mode signaling is promising to remove this bottleneck.
- Through simulation, circuit fabrication and actual measurements, we have demonstrated that current mode signaling has overwhelming advantages over the currently used voltage mode buffer insertion schemes.
- We have demonstrated that the particular configuration suggested by us for a current mode scheme is superior to other current mode schemes.
- Our scheme is robust with respect to batch to batch parametric variations and to on chip parametric variation.
- Therefore we assert that it is a practical option for use in modern systems for implementing both unidirectional and bidirectional data links.