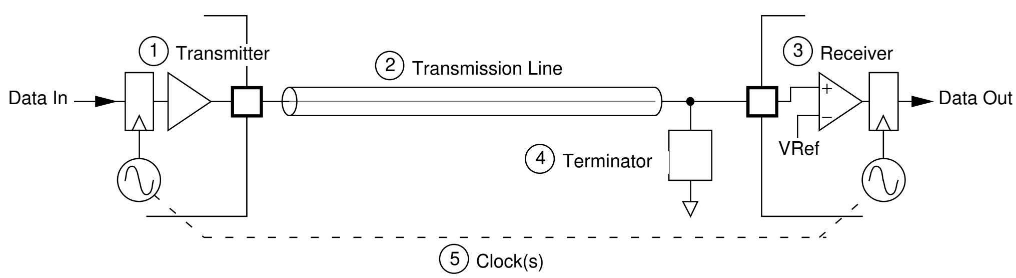
Signaling in High-Performance Memory Systems

John Poulton

Research Professor
Department of Computer Science
University of North Carolina at
Chapel Hill

jp@cs.unc.edu
http://www.cs.unc.edu/~jp

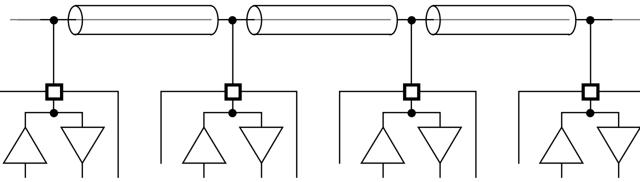
What's a Signaling System?



1. **Transmitter** encodes data as voltage/current signal levels onto the line
2. **Transmission line** delivers the signal to the receiver
3. **Receiver** compares signal with reference to recover data
4. **Terminator** removes signals from line, once they're received
5. **Clocks** tell transmitter when to drive a new signal, receiver when to sample

For this discussion, we'll assume

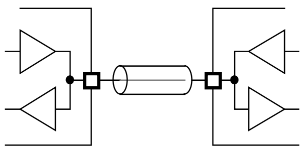
- binary (two-level) signaling
- direct encoding (*e.g.* V_{Hi} = "1", V_{Lo} = "0")



A **Bus** has
- multiple transmitters and receivers ("drops"), often paired up as transceivers
- transceivers share a common transmission line
- any one drop may send; any (or all) drops receive at a given time

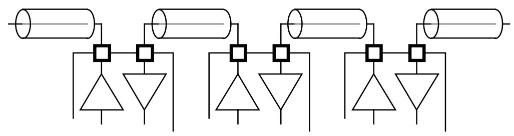
---

This is **not** a bus:



...it's a bidirectional link.

This isn't a bus, either:



...it's a linear network.

Outline

5–28
- Transmission Lines
  - ° Impedance, propagation, reflection
  - ° Impedance discontinuities
  - ° Termination methods
  - ° Loss mechanisms
  - ° Cross-talk, modes, impedance matrices
  - ° Typical properties of PC-board T-lines

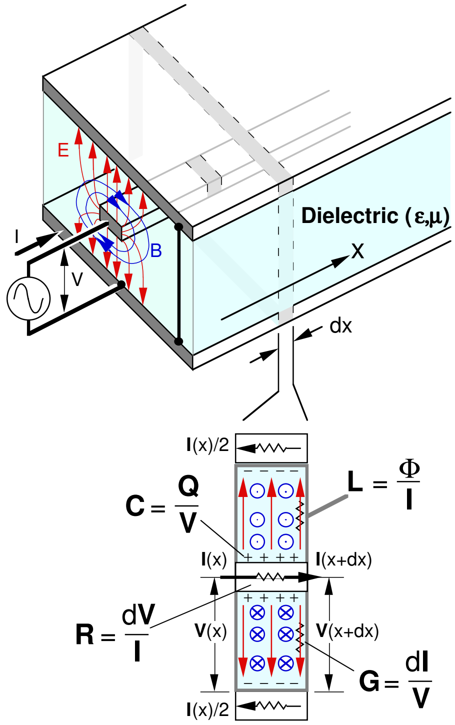
29–35
- Signaling Methods
  - ° Voltage-mode vs Current-mode
  - ° Single-ended vs Differential
  - ° References
  - ° Clocking

36–55
- Bus-based Signaling Systems
  - ° Problems: lowered Z_0, v, f_c
  - ° CMOS , voltage-mode signaling
  - ° SSTL & SLDRAM
  - ° Open-drain systems, BTL, GTL, etc
  - ° RAMBUS
  - ° Summary

56–57
- Alternative Signaling Approaches
  - ° Future Prospects
  - ° Serial Links

- Wrapup, Questions, Discussion

Transmission Lines



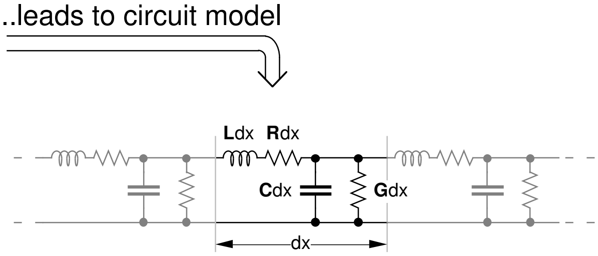
If I,V change at the drive point, **B**,**E** change as well; disturbance propagates away from the drive point at the speed of light

$$v = \frac{1}{\sqrt{\varepsilon\,\mu}} \approx \frac{c}{\sqrt{\varepsilon_r}}$$

IF the dielectric is homogeneous and $\mu = \mu_0$ (as usual)

(Quasi-TEM guided wave propagation)

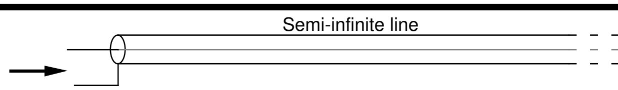
...leads to circuit model



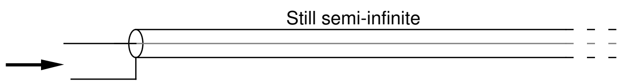
L,R,C,G are per unit length

Characteristic Impedance

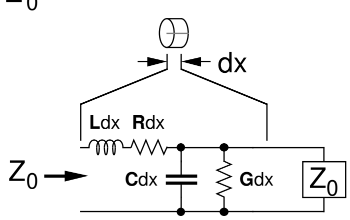
Impedance is some value Z_0



Peel off dx worth of line; impedance still Z_0



So, original line looks like this:



$$Z_0 = Rdx + j\omega Ldx + \frac{1}{j\omega Cdx + Gdx + 1/Z_0}$$

$$(Z_0 - Rdx - j\omega Ldx)(j\omega Cdx + Gdx + 1/Z_0) = 1$$

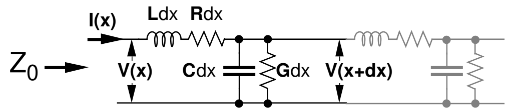
$$Z_0(G+j\omega C)dx - \cancel{(R+j\omega L)dx(G+j\omega C)dx} - \frac{(R+j\omega L)dx}{Z_0} = 0$$

$$\boxed{Z_0 = \sqrt{\frac{R+j\omega L}{G+j\omega C}} = \sqrt{\frac{L}{C}} \text{ when } R=G=0}$$

Impedance of a (near-) lossless line is frequency independent and real--just a resistance!

Propagation Constant

$$\frac{\partial V(x)}{\partial x} = -(R+j\omega L)I(x) = \frac{-(R+j\omega L)V(x)}{Z_0}$$



$$\frac{\partial V(x)}{\partial x} = -[(G+j\omega C)(R+j\omega L)]^{1/2}\, V(x)$$

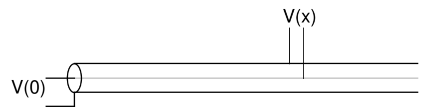
$$V(x) = V(0)\, e^{-Ax}\ ;\quad A = [(G+j\omega C)(R+j\omega L)]^{1/2}$$

$$A = (j\omega RC + j\omega GL - \omega^2 LC)^{1/2}$$

$$= j\omega\sqrt{LC}\left(1-j\frac{RC+GL}{\omega LC}\right)^{1/2} \text{ Taylor exp.}$$

$$= j\omega\sqrt{LC}\left(1-j\frac{RC+GL}{2\omega LC}\right)$$

$$= j\omega\sqrt{LC} + \frac{R}{2Z_0} + \frac{GZ_0}{2}$$

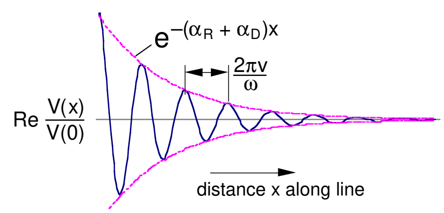


Phase change along line, where wave velocity is $v = \dfrac{1}{\sqrt{LC}} = \dfrac{1}{\sqrt{\varepsilon\,\mu}}$

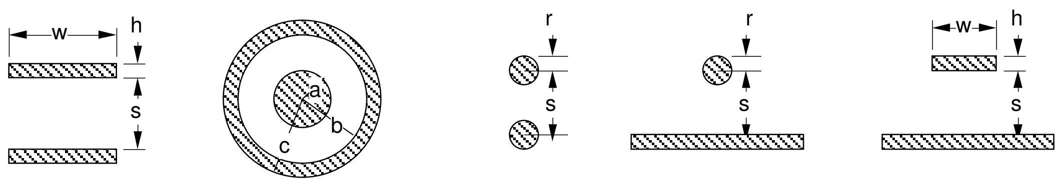
$$(x) = V(0)\, e^{-j\omega\sqrt{LC}x}\ \underbrace{e^{-(\alpha_R + \alpha_D)x}}_{\textbf{Attenuation}}$$

$$\alpha_R = \frac{R}{2Z_0} \qquad \alpha_D = \frac{GZ_0}{2}$$

conductor loss **dielectric loss**



Characteristic Impedance of Some Typical Lines



$R_{DC} = \dfrac{2\rho}{wh}$	$R_{DC} = \dfrac{\rho}{\pi a^2} + \dfrac{\rho}{\pi(c^2 - b^2)}$	$R_{DC} = \dfrac{2\rho}{\pi r^2}$	$R_{DC} = \dfrac{\rho}{\pi r^2}$	$R_{DC} = \dfrac{\rho}{wh}$
$C = \dfrac{\varepsilon w}{s}$	$C = \dfrac{2\pi\varepsilon}{\log(b/a)}$	$C = \dfrac{\pi\varepsilon}{\log(s/r)}$	$C = \dfrac{2\pi\varepsilon}{\log(2s/r)}$	$C \approx \dfrac{\varepsilon w}{s} + \dfrac{2\pi\varepsilon}{\log(4s/h)}$
$Z_0 = \sqrt{\dfrac{\mu}{\varepsilon}}\,\dfrac{s}{w}$	$Z_0 = \sqrt{\dfrac{\mu}{\varepsilon}}\,\dfrac{\log(b/a)}{2\pi}$	$Z_0 = \sqrt{\dfrac{\mu}{\varepsilon}}\,\dfrac{\log(s/r)}{\pi}$	$Z_0 = \sqrt{\dfrac{\mu}{\varepsilon}}\,\dfrac{\log(2s/r)}{2\pi}$	$Z_0 = \dfrac{\sqrt{\varepsilon\mu}}{C}$

Good only for homogeneous dielectrics, where $L = \varepsilon\mu/C$

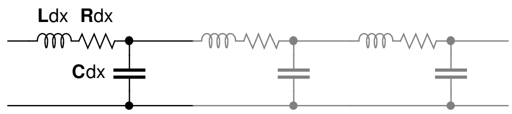
Caveats:

1. Approximations only, useful for 1rst-order estimates; do **NOT** use these formulas for anything serious! Fire up your field solver...
2. Good only for homogeneous dielectrics; PC-board microstrips are not homogeneous.
3. Since $\mu = \mu_0$ for many dielectrics,

$$\sqrt{\frac{\mu}{\varepsilon}} = \sqrt{\frac{\mu_0}{\varepsilon_0}} \bullet \frac{1}{\sqrt{\varepsilon_r}} = \frac{377\Omega}{\sqrt{\varepsilon_r}}$$

Main thing to note: Z_0 is always a function of **ratios** of dimensions!

RC vs LC Behavior



In many situations, it's reasonable to set G=0 (dielectrics are pretty good for many signaling systems)

What about R?

If $j\omega L << R$, then the line is mainly an RC transmission line

If $j\omega L >> R$, then mainly an LC line

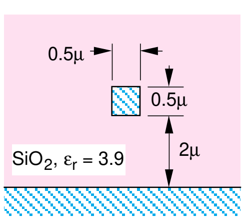
Every line has some frequency

$$f_0 = \frac{R}{2\pi L}$$

below which it's RC and above which it's LC

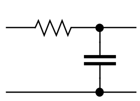
RC & LC Approximations

A typical on-chip wire:

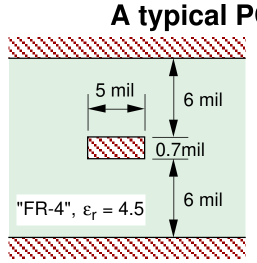


$L = 0.6$ nH/mm
$C = 73$ fF/mm
$R_{dc} = 120\ \Omega$/mm
$f_0 = 32$ GHz

The RC model works for CMOS!
(...well, mostly...)

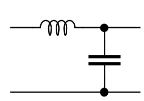


A typical PC-board wire:



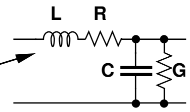
$L = 0.5$ nH/mm
$C = 104$ fF/mm
$R_{dc} = 0.008\ \Omega$/mm
$f_0 = 2.5$ MHz

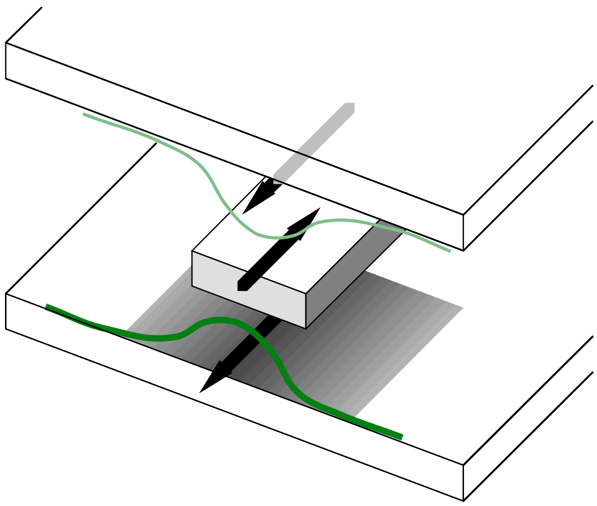
The LC model is OK for PC board traces
(mostly... $R = R(f)$ from skin effect)



Signal Returns

Impedance of line
AND it's return!





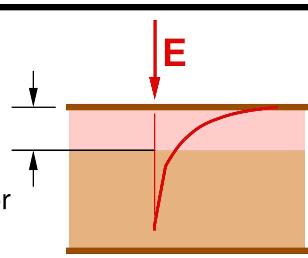
In striplines, return current in planes is concentrated near signal conductor, so return impedance is significant

Skin Effect

Skin depth:

$$\delta = \sqrt{\frac{\rho}{\pi f \mu}}$$

depth into conductor where E falls to 1/e of surface value

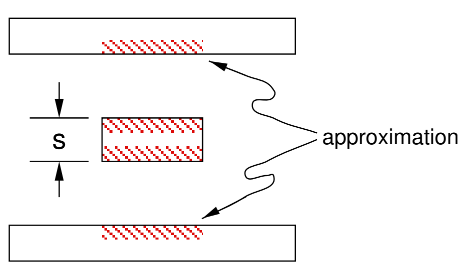


Skin-depth frequency:

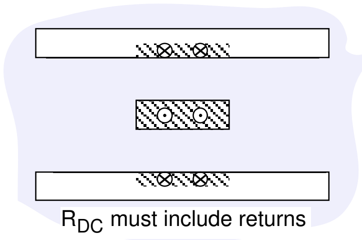
$$f_s = \frac{4\rho}{\pi \mu s^2} = \frac{4\rho}{\pi \mu_0 s^2}$$

non-magnetic conductors

frequency at which skin-depth equals half the conductor thickness



Conductor Loss



$$R(f) = R_{DC}\sqrt{\frac{f}{f_s}}$$

$$\alpha_C(f) = \frac{DC}{2Z_0}\sqrt{\frac{f}{f_s}}$$

$$\frac{(x)}{V(0)} = e^{-\alpha_C L}$$

$$\text{Attn (dB/m)} = 10\log_{10}\left(\frac{P(1m)}{P(0)}\right)$$

$$= 20\log_{10}\left(\frac{V(1m)}{V(0)}\right)$$

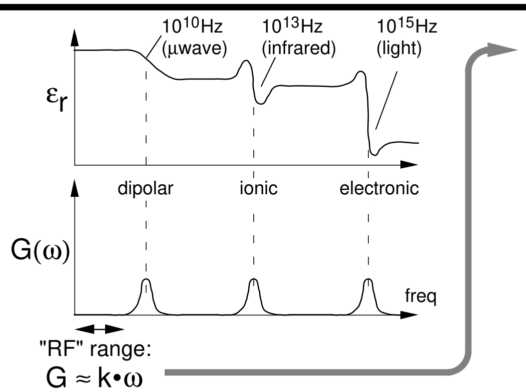
$$= 8.686 \bullet \alpha_C$$

Some typical signal conductors		
Conductor	f_s	**Attn (@ 1GHz)**
RG-223/U	**22.5 KHz**	**0.5 dB/m**
5-mil, 0.5 Oz PCB trace	**56 MHz**	**6 dB/m**
0.5µ x 0.5µ on-chip wire	**113 GHz**	**same as DC!**

$$\frac{V(6")}{V(0)} = 0.90 \text{ @ 1GHz}$$

possibly a concern for bus-based memory signaling

Dielectric Loss



...so, can define

$$\tan\delta = \frac{G}{\omega C}$$

$$\alpha_D = \frac{GZ_0}{2} = \frac{2\pi f C \tan\delta \sqrt{L/C}}{2}$$

$$= \pi f \tan\delta \sqrt{LC}$$

$$\boxed{\alpha_D = \frac{\pi\sqrt{\varepsilon_r}\, f \tan\delta}{c}}$$

only for homogeneous dielectrics!

Some typical PC Board Dielectrics				
Dielectric		ε_r	**tanδ**	**8.86α_D @1GHz**
"FR-4"	**$**	**4.7**	**0.035**	**7.0 dB/m**
GETEK	**$.$**	**3.9**	**0.010**	**1.8 dB/m**
Teflon	**$$$**	**2.25**	**0.001**	**0.14 dB/m**

$$\frac{V(6")}{V(0)} = 0.89 \text{ @ 1GHz}$$

so, possibly a concern for memory interconnect

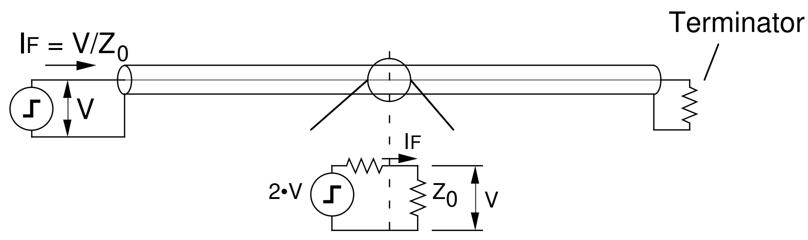
Terminators

Real lines (PCB traces, cables) are fairly low-loss

- Good news: signals get from here to there with little attenuation
- Bad news: if the signals aren't removed, they'll bounce around and interfere with later signals (a.k.a, **inter-symbol interference**).

The **terminator** gets rid of the signal once it's received

To see what happens at terminations (and other discontinuities), useful to introduce a mythical voltage generator that represents the traveling wave.



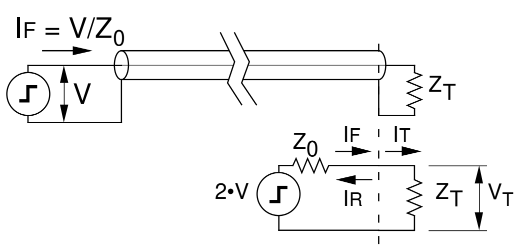
General Terminations

Mythical voltage divides across line & terminator → $V_T = 2V \bullet \frac{Z_T}{Z_0 + Z_T}$

producing terminator current... → $I_T = \frac{V_T}{Z_T} = \frac{2V}{Z_0 + Z_T}$

Σ currents into output node = 0 → $I_R = I_F - I_T = \frac{V}{Z_0} - I_T$

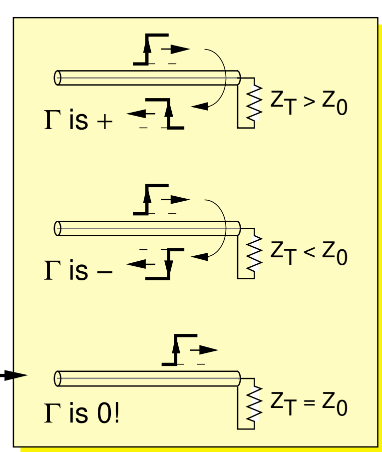
$$= \frac{V(Z_T - Z_0)}{Z_0(Z_T + Z_0)}$$



Reflection Coefficient (Γ)

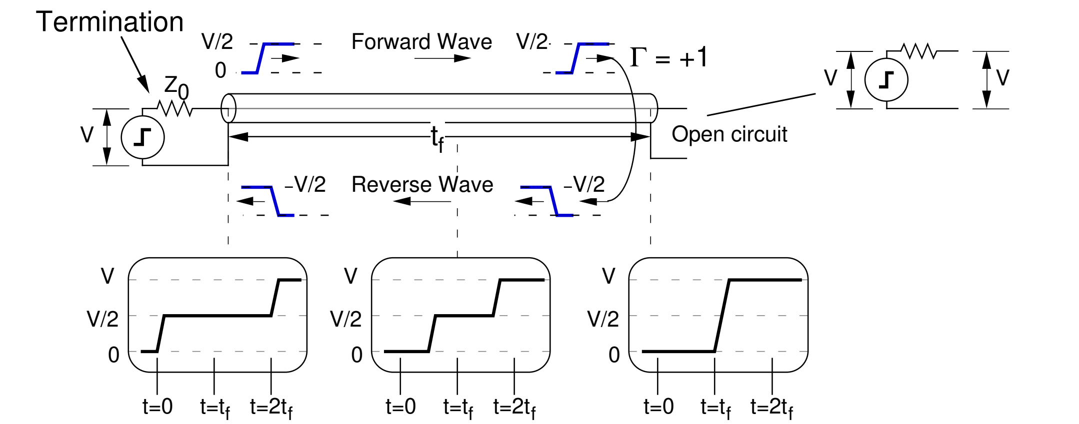
$$\boxed{\Gamma = \frac{I_R}{I_F} = \frac{V_R}{V_F} = \frac{Z_T - Z_0}{Z_T + Z_0}}$$

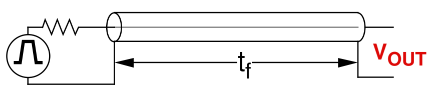
This situation is highly desirable; forward wave is completely absorbed at the receiver.



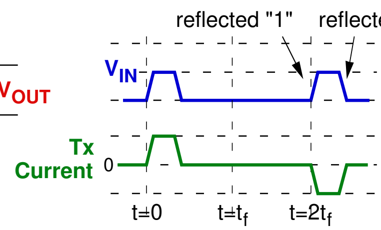
Source Termination

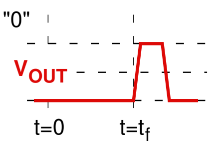
Can just as well put the terminator at the transmitter (source, or series termination)





What happens when pulse width < t_f ??

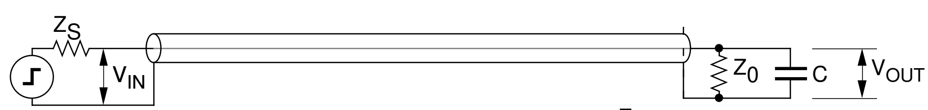




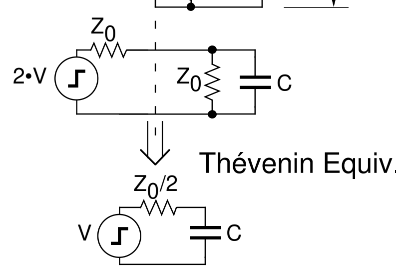
Current only flows when tx opposes reflected wave, so low power!

Imperfect Terminations: Capacitance

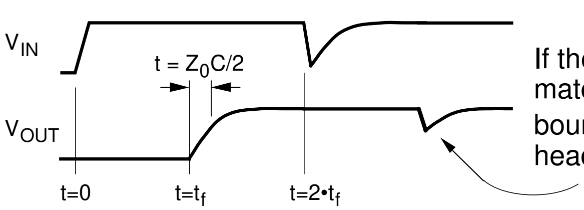
Suppose your 'perfect' terminator is shunted by a capacitor (as it usually is!!):



Bandwidth is limited by cutoff frequency of Thévenin network, $f_0 = \frac{1}{\pi Z_0 C}$



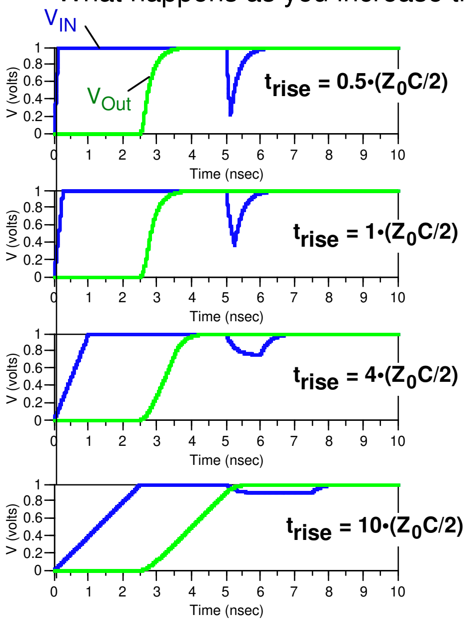
If the forward wave is a fast-rise edge, the capacitor generates a large reflection:

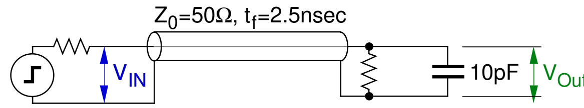


If the transmitter isn't perfectly matched ($Z_S \neq Z_0$), the reflection bounces off the transmitter and heads toward the receiver again.

Managing Reflections with Capacitive Termination

What happens as you increase the risetime??





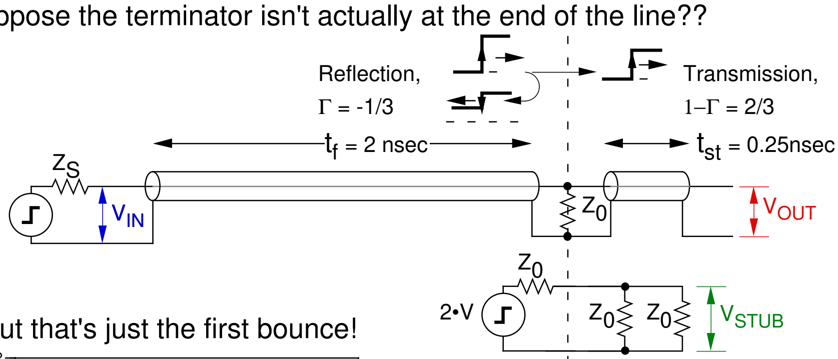
As a rule of thumb, if you want reflection to be (1/n)th of signal, risetime should be $n\bullet Z_0C/2$.

...or, need to terminate the transmitter as well as the receiver, to absorb reflections.

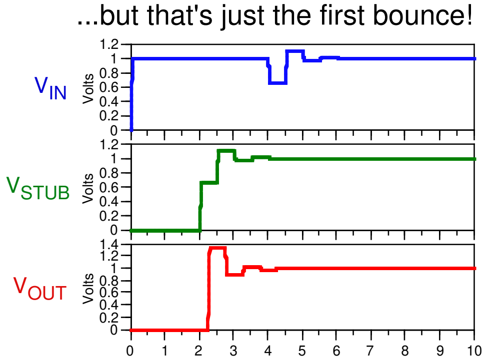
Series termination works fine (no terminator at rx), but bandwidth will be halved, *i.e.*, $1/(2\pi Z_0C)$ instead of $1/(\pi Z_0C)$

Imperfect Terminations: Stubs

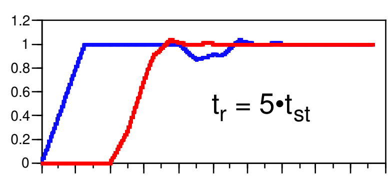
Suppose the terminator isn't actually at the end of the line??



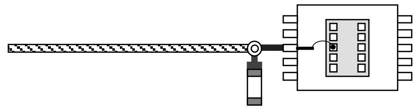
...but that's just the first bounce!

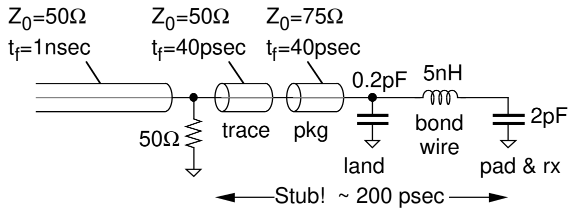


As with cap's, can mitigate the evil effects of stubs with risetime control.



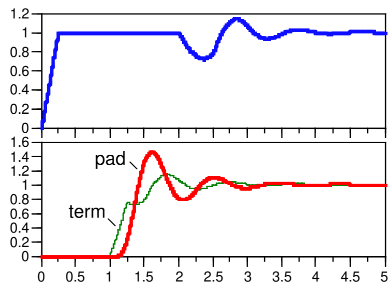
A Typical Stub Termination





Signal at Tx (0.25nsec edge)

Internal signal at Rx pad & at terminator

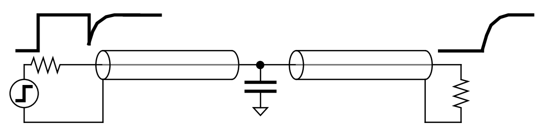


Generally, you'd probably like to avoid this type of behavior.

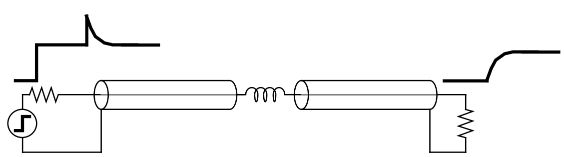
- Edge-rate control
- Careful board layout
- Advanced packaging
- On-chip termination

T-Line Discontinuities

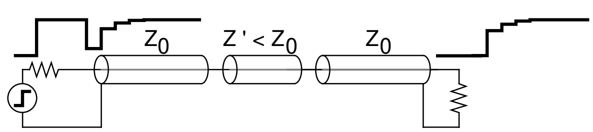
Capacitive discontinuity:
works pretty much like the capacitively shunted terminator



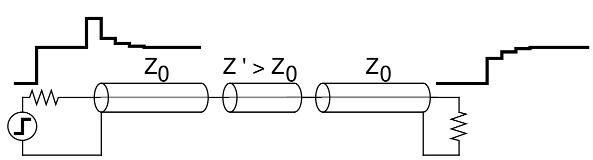
Inductive discontinuity:
the complementary behavior to the capacitive one



Impedance discontinuity:
with $Z' < Z_0$

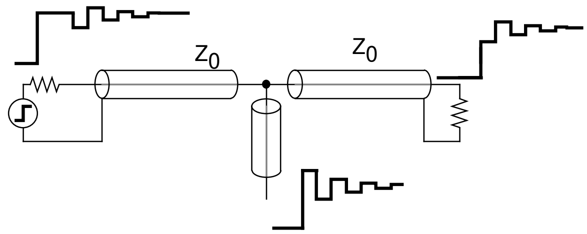


Impedance discontinuity:
with $Z' > Z_0$

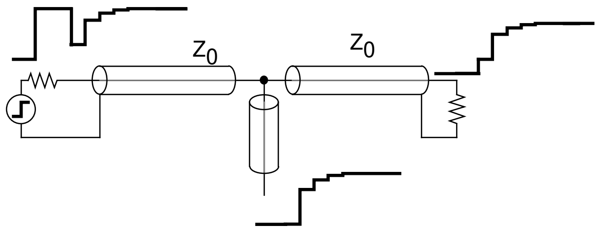


One thing they have in common: they're **all** low-pass!

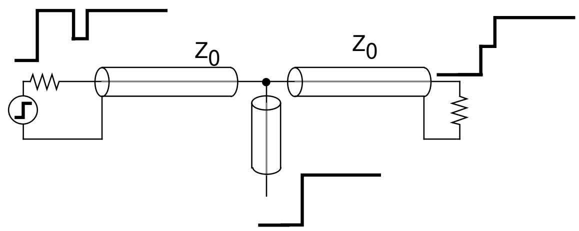
Stub Discontinuities



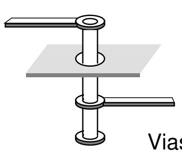
Open Stub
with $Z' < Z_0$

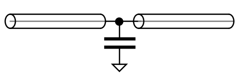


Open Stub
with $Z' = 0.5 \cdot Z_0$

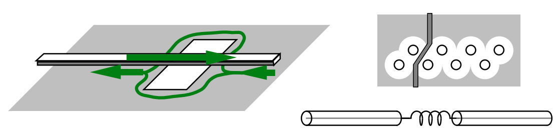


Some Common Discontinuities

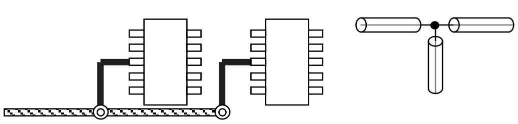




Vias are small capacitive discontinuities (C ~ 0.1pF typically)

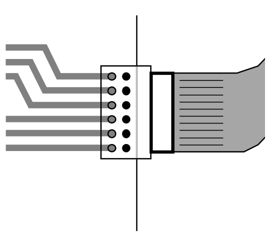


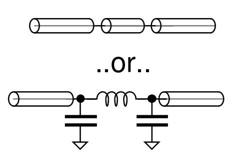
Ground-plane slots are inductive discontinuities (also pretty good antennas!). Often created unintentionally in dense fields of vias.



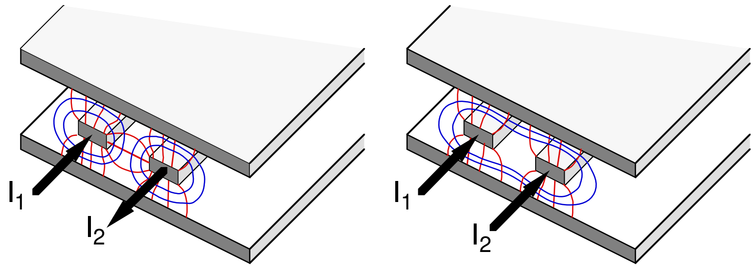
Multi-drop signals introduced stub discontinuities.

Trace squeezes introduce in-line low-impedance section.





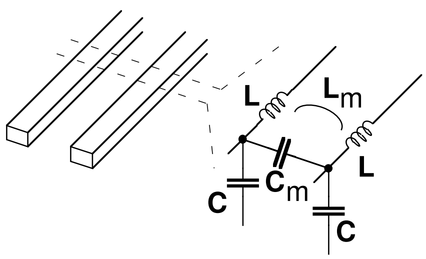
Connectors introduce impedance discontinuities, often modeled with a combination of lumped elements.

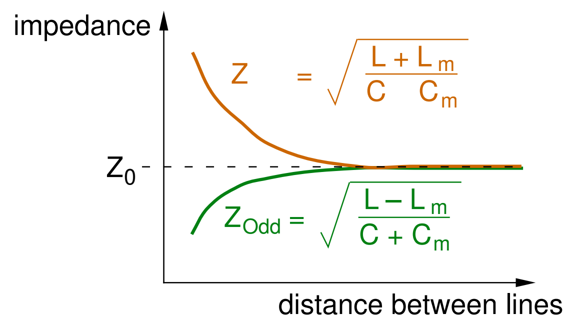


"Odd-mode" excitation

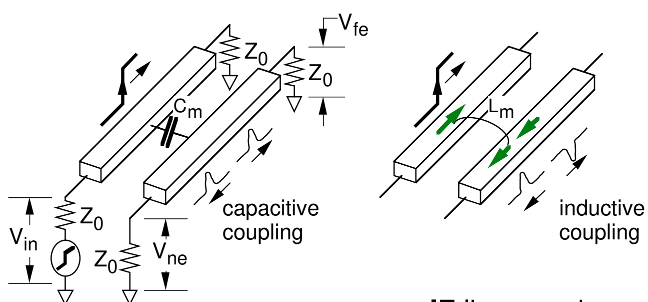
Field distribution depends on relative directions of currents in neighboring lines...

...and so does impedance!

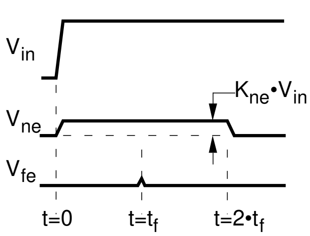




Crosstalk



IF lines are homogeneous
AND all ports are terminated...

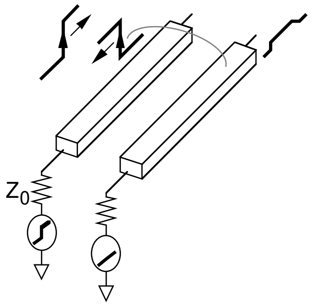


equal, when homogeneous

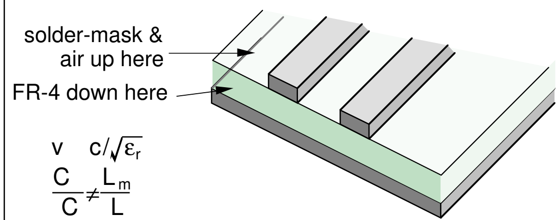
$$K_{ne} = \frac{V_{ne}}{V_{in}} = \frac{1}{4}\left(\frac{C_m}{C} + \frac{L_m}{L}\right) = \frac{C_m}{2C}$$

$$K_{fe} = \frac{V_{fe}}{V_{in}} = \frac{1}{4}\left(\frac{C_m}{C} - \frac{L_m}{L}\right) = 0$$

In inhomogeneous lines, far-end crosstalk gets worse with
- faster edges
- longer lines



Series-terminated back-wave induces far-end crosstalk on neighbor!

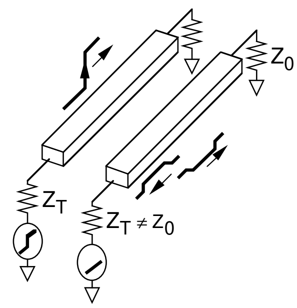


$v \;\; c/\sqrt{\varepsilon_r}$

$\frac{C}{C} \neq \frac{L_m}{L}$

Microstrips aren't homogeneous; watch out for far-end crosstalk.

Near-end crosstalk can become a far-end problem if transmitters aren't terminated.



Multiconductor Transmission Lines

For 2 lines, the **Telegrapher's Equations** are

$$\frac{d}{dx}\begin{bmatrix} V_1(x) \\ V_2(x) \end{bmatrix} = -\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \bullet \begin{bmatrix} I_1(x) \\ I_2(x) \end{bmatrix} - j\omega \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \bullet \begin{bmatrix} I_1(x) \\ I_2(x) \end{bmatrix}$$

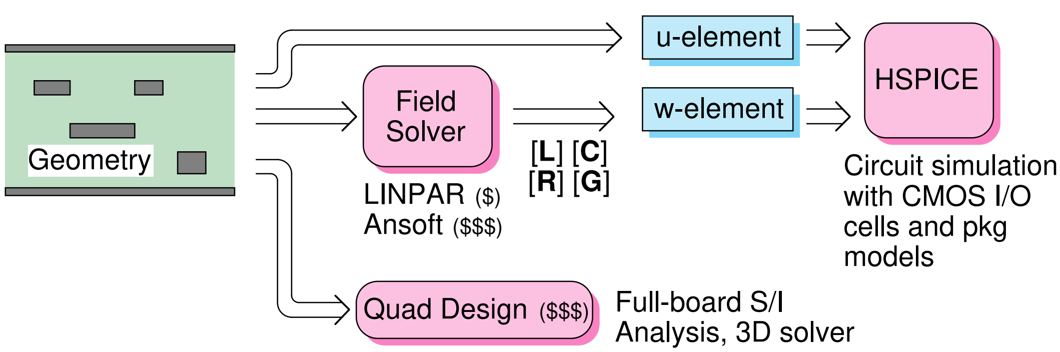
$$\frac{d}{dx}\begin{bmatrix} I_1(x) \\ I_2(x) \end{bmatrix} = -\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \bullet \begin{bmatrix} V_1(x) \\ V_2(x) \end{bmatrix} - j\omega \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \bullet \begin{bmatrix} V_1(x) \\ V_2(x) \end{bmatrix}$$

These generalize for N coupled lines to

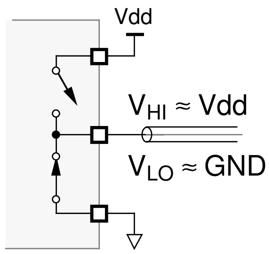
$$\frac{d}{dx}[\mathbf{V}] = -[\mathbf{R}]\cdot[\mathbf{I}] - j\omega[\mathbf{L}]\cdot[\mathbf{I}]$$

$$\frac{d}{dx}[\mathbf{I}] = -[\mathbf{G}]\cdot[\mathbf{V}] - j\omega[\mathbf{C}]\cdot[\mathbf{V}]$$

where [**L**], [**C**], [**R**], and [**G**] are NxN matrices

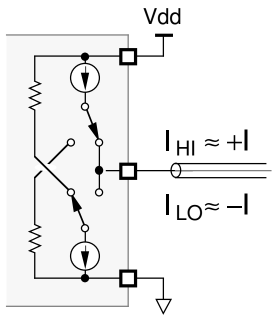


Signaling - Voltage-Mode vs Current-Mode



Voltage-Mode Signaling

"1" and "0" represented by voltages, usually the two power supplies. Low impedance switches required to get line voltages near power supply voltages; these generate the dreaded "simultaneous switching" noise.



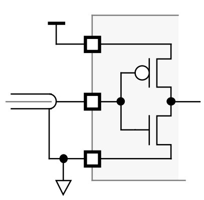
Current-Mode Signaling

"1" and "0" represented by +/– current.

In well-designed I-mode signaling, I_{PS} never changes. High-impedance current sources isolate the line from the power supplies.

Signal is converted to a voltage in the receiver's terminator.

Signaling - References



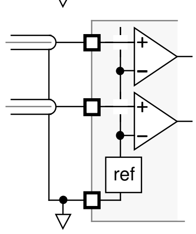
Built-in Reference

PFET/NFET split power supply voltage to get reference.

$$V_{inv} = \frac{V_{TN} + \sqrt{\beta_P/\beta_N}\,(V_{DD}+V_{TP})}{1 + \sqrt{\beta_P/\beta_N}}$$

Process dependent

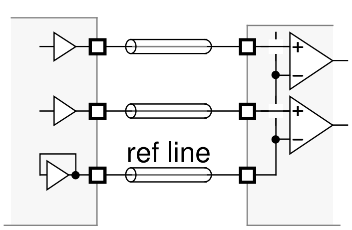
Whims of power supply gods



Local Precision Reference

External or internal; usually shared among many receivers.

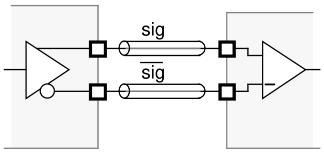
Reference fixed, more or less, but noise across signal return impedance isn't cancelled...



Transmitter-generated (Bundled) Reference

Hope that noise injected into signal lines is tracked by reference line; unfortunately, Ref Line sees different impedance at far end, since shared by many Rx's.

Differential Signaling



Get rid of pesky reference problem by sending a "reference" (actually $\overline{\textbf{signal}}$) along with every signal.

GOOD:

Almost all sources of noise are now common-mode, so can get by with tiny signal voltages, low power

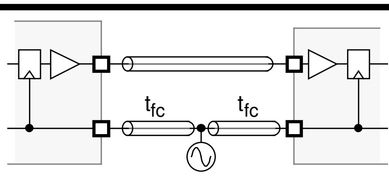
BAD:

Two wires required, instead of one. Need to keep wires close together, tightly coupled, same length; sometimes an "interesting" engineering problem.

Unproved (but widely believed) Lemma:

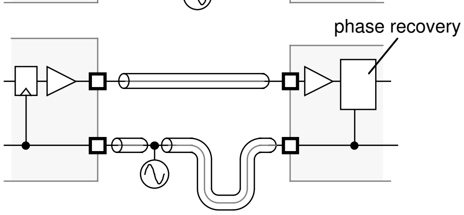
It is always possible to operate a differential signaling system >2x as fast as a single-ended one, all other things being equal.

Clocking



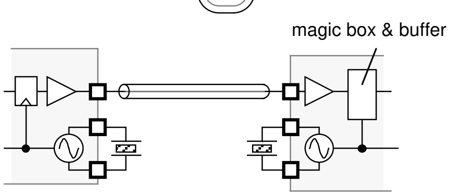
Synchronous

Every participant gets same frequency and phase.



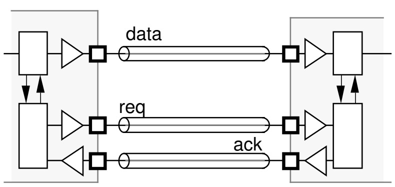
Mesochronous

Every participant gets same frequency, but unknown phase. Requires a way to recover the phase from the data. Coding (*e.g.* 8b/10b) is often used to make sure there are sufficient data edges.



Plesiochronous

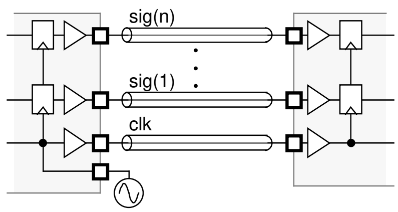
Every participant gets nearly the same frequency, slowly drifiting phase. Requires a way to detect when the Rx clock has drifted 1/2 cycle from Tx clock.



Asynchronous

Dispense with clocks altogether, use (*e.g.*) request/acknowledge 4-phase handshake to ensure correct sequencing of events.

Clock Forwarding



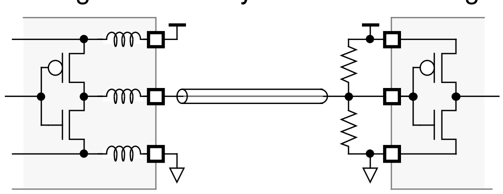
Send a clock along with (a bundle of) data; clock and data "enjoy" the same delay and impedance discontinuties, so skew between data and clock is minimized.

Currently very fashionable, and works quite well up to ~ 1GHz.

At higher frequencies, gets harder to keep the delays matched sufficiently well to maintain synchronous operation.

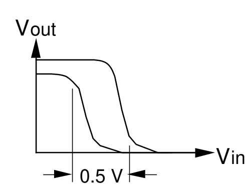
Why Worry About Signaling?

So, what's wrong with the way we've been doing things for years?



Reference

Inverters make really terrible receivers. Often as much as 500 mV gap between guaranteed HI and LO.

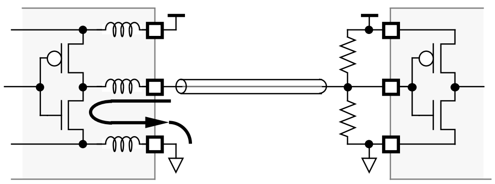


Termination

Termination (for minimum tx current) needs to be at about Vinv, requiring a second power supply or voltage divider. In 50Ω environment, divider consumes ~50 mW even without a signal!

Transmitter FETs biased into linear region, so low resistance; any reflections from impedance discontinuities bounce off the transmitter and head for the receiver.

Signaling? ...cont.



Need ± 1 volt of signal, so ± 20 mA of signal current.

$$V_{bond} = L_{bond}\frac{dI}{dt} = 5nH \bullet \frac{20mA}{0.1nsec} = 1 \text{ volt} \quad \text{Yikes!}$$

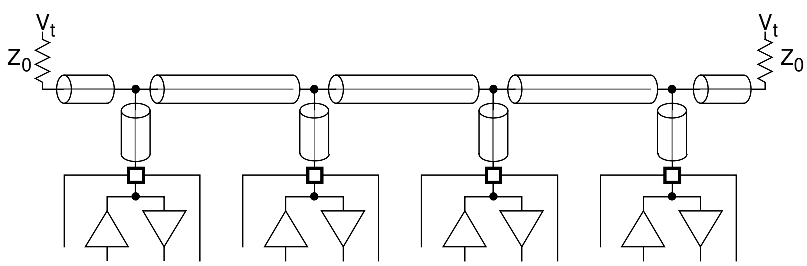
...and several Tx's may have to share a common supply lead...

Voltage-mode signaling systems themselves generate most of the noise that they must overcome. $V_{Noise} \propto V_{Signal}$, a loosing proposition.

Power

~10-20 mW dissipated in the transmitter, and ~70 mW in the terminator.
Does it really take nearly 100 mW to send a bit from one chip to another??

What's a Bus?



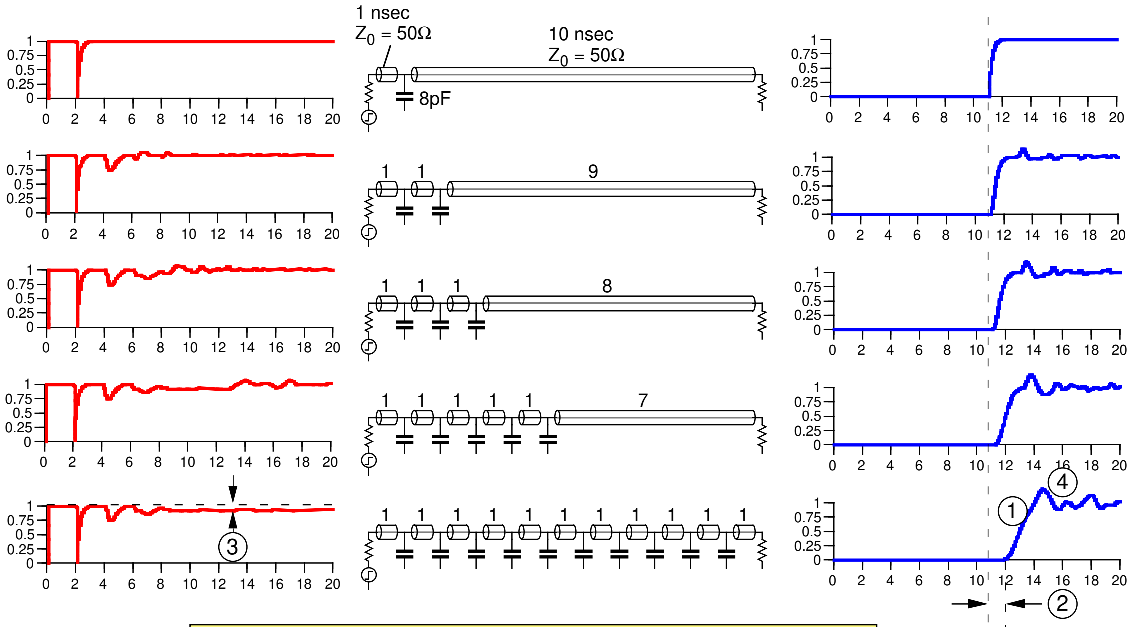
It's a transmission line with multiple discontinuities.

Each driver sees a line impedance of $Z_0/2$,

...so, to maintain receiver voltage, need twice the driver power as in a point-to-point signaling system.

This is kinda bad, but what about those discontinuities??

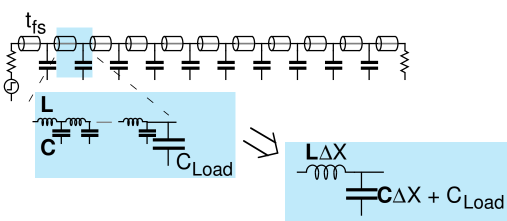
Effects of Multiple Capacitive Point Discontinuities



1. Reduced bandwidth
2. Increased delay
3. Reduced impedance
4. Overshoot, ringing

Lumped Model

if $t_{rise} >> 2 \bullet t_{fs}$ ...



$$Z_{Loaded} = \sqrt{\frac{\mathbf{L}\Delta x}{\mathbf{C}\Delta x + C_{Load}}} = \frac{Z_0}{\sqrt{1 + C_{Load}/\mathbf{C}\Delta X}}$$

So, define "bus factor" B_f

$$B_f = \sqrt{1 + C_{Load}/\mathbf{C}\Delta x} = \sqrt{1 + C_{Load}/C_{Seg}}$$

Charateristic impedance $Z_{Bus} = \dfrac{Z_0}{B_f}$

Wave velocity $v_{Bus} = \dfrac{v_{Line}}{B_f}$

Cutoff frequency $F_0 = \dfrac{1}{\pi\, t_{fs} B_f}$

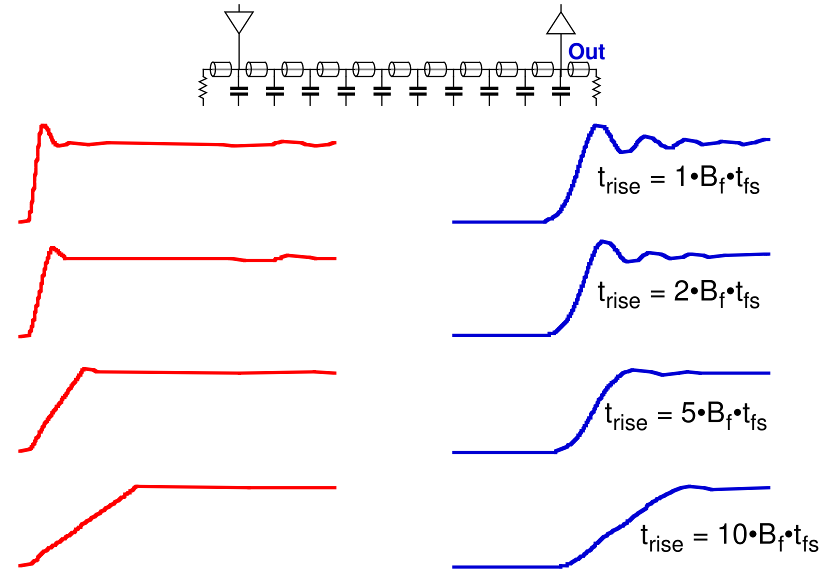
Time of flight through a bus segment $T_0 = B_f\, t_{fs}$

Cutoff frequency for a lumped-element line is:

$$F_0 = \frac{1}{\pi\sqrt{LC}} = \frac{1}{\pi\sqrt{\mathbf{L}\Delta x \bullet (\mathbf{C}\Delta x + C_{LOAD})}} = \frac{1}{\pi \Delta x \sqrt{\mathbf{LC}}\, B_f} = \frac{1}{\pi\, t_{fs} B_f}$$

Nice discussion in Tom Lee's *The Design of CMOS Radio-Frequency Integrated Circuits* ... a terrific book!

Allowable Risetime

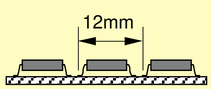
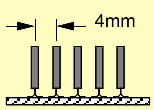
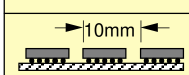


Rule of thumb: $t_{rise} \geq 5 \bullet B_f \bullet t_{fs}$ to reduce overshoot & ringing

and... $F_{Max} = \dfrac{1}{4 \bullet t_{rise}} = \dfrac{1}{4 \bullet 5 \bullet T_0} \approx 0.15 \bullet F_0$

Some Examples

Assumes $Z_0 = 70\Omega$, $1/v = 7.2$ psec/mm

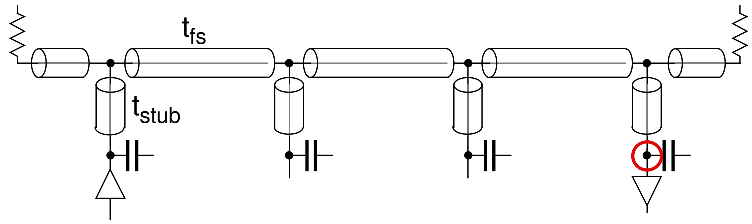
Memory		Pkg	C_{Load}	C_{Seg}	B_f	Z_{Bus}	$t_{drop\text{-}to\text{-}drop}$	F_0
	SDRAM	400-mil TSOP	6.5 pF	1.25 pF	2.49	28 Ω	215 psec	1.48 GHz
	SLDRAM (SLIO)	VSMP	3 pF	0.42 pF	2.85	25 Ω	82 psec	3.88 GHz
	SLDRAM (LVCMOS)		6.5 pF		4.06	17 Ω	116 psec	2.72 GHz
	RDRAM	CSP	2.4 pF	1.0 pF	1.84	38 Ω	132 psec	2.40 GHz

Observations:

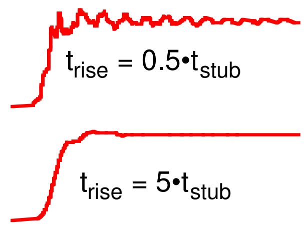
Try to increase F_0 by making t_{fs} smaller, then Z_{Bus} gets smaller too.

Try to increase Z_{Bus} by making Z_0 higher, then F_0 decreases.

Stubs



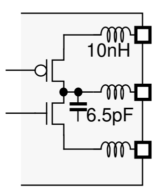
$$t_{rise} \geq \begin{cases} 5 \bullet t_{fs} \sqrt{1 + (C_{Load} + C_{stub}) / \mathbf{C}\Delta x} \\ t_{stub} \end{cases}$$



Lines with stubs are complex circuits.

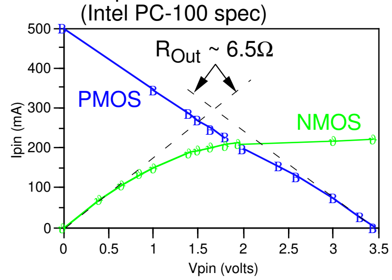
Accurate modeling and circuit simulation required!

Voltage-Mode Bus Signaling – PC-100

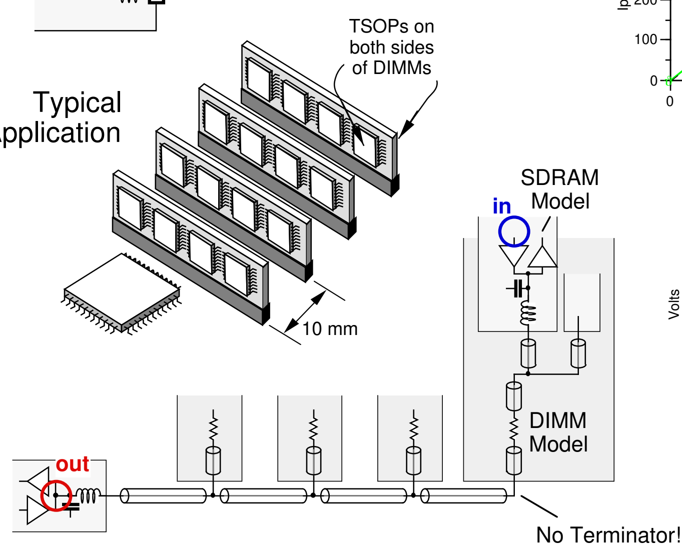


SDRAMs: Power pin per driver (x4's & x8's); shared on x16's.

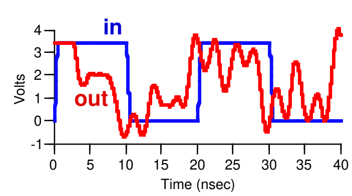
Output Characteristics (Intel PC-100 spec)



Typical Application

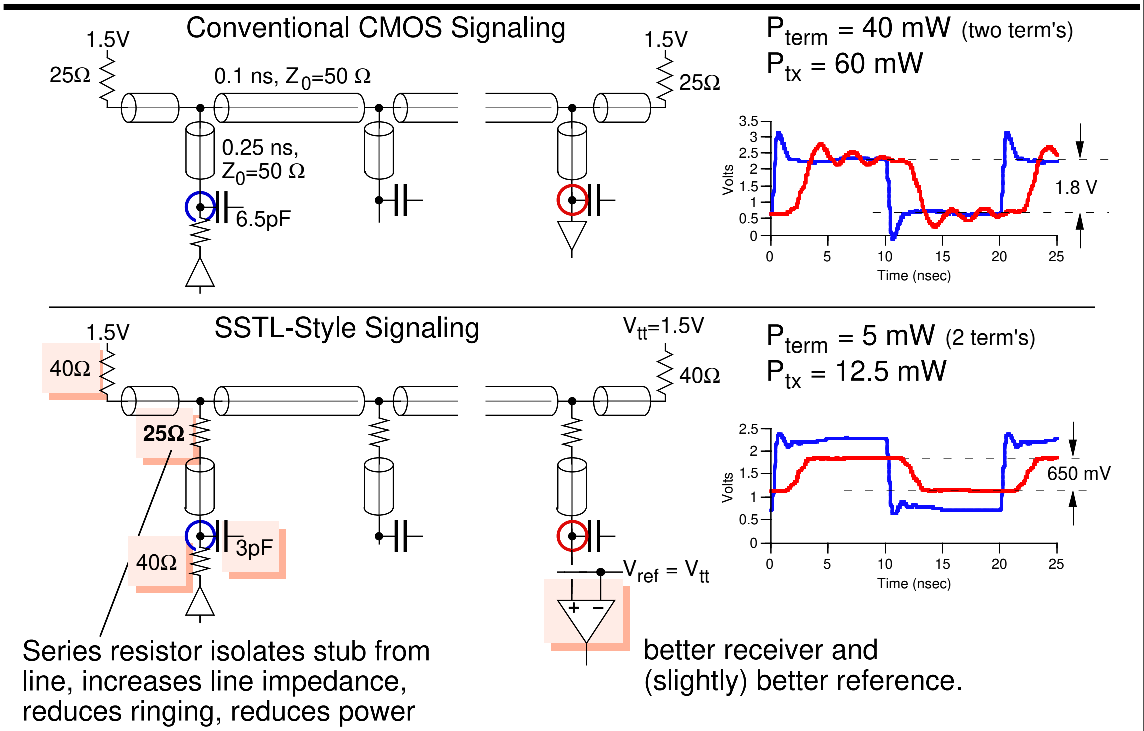


Simulation of simple model:

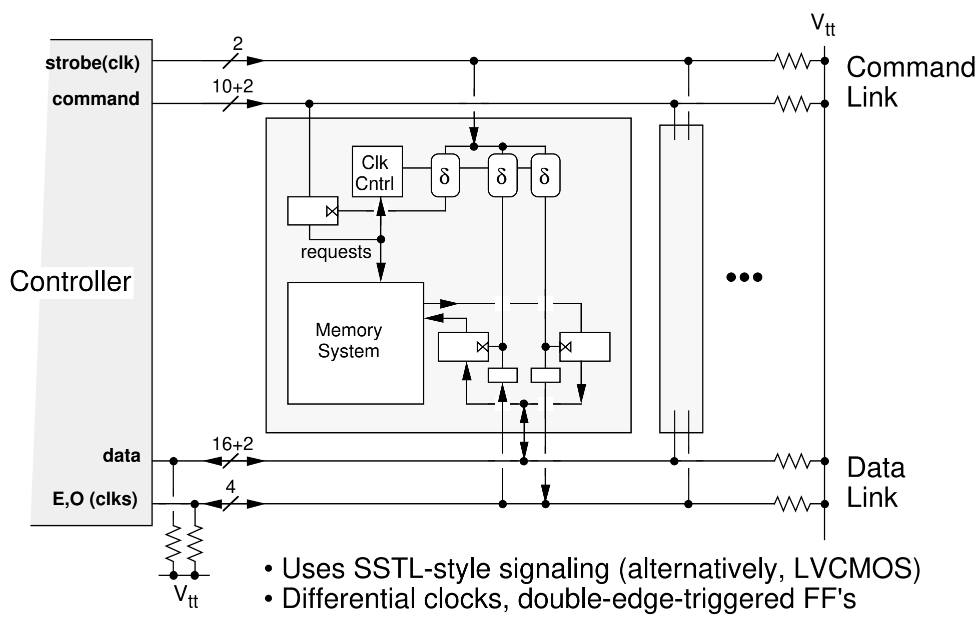


Even without termination, P_d ~ 100 mW/pin

Series Stub Terminated Logic (SSTL)



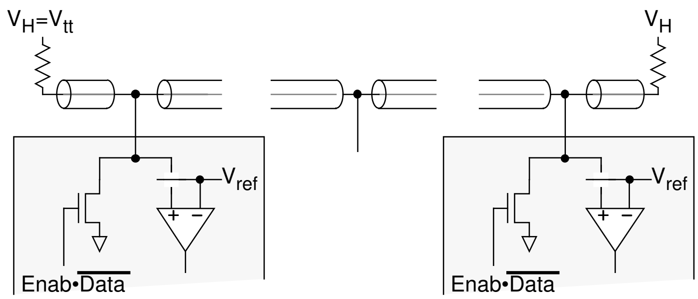
SLDRAM (SyncLink DRAM)



- Uses SSTL-style signaling (alternatively, LVCMOS)
- Differential clocks, double-edge-triggered FF's
- Source-synchronous timing (controller-adjustable δ's)
- 200 MHz (800 MByte/sec) operation

Perhaps more relevantly: DDR's use SSTL!

Wired-OR Bus Signaling



In wired-OR scheme:

Driver supplies the "0's"
Terminator supplies "1's"

Export power dissipation from on-chip driver to terminator resistor

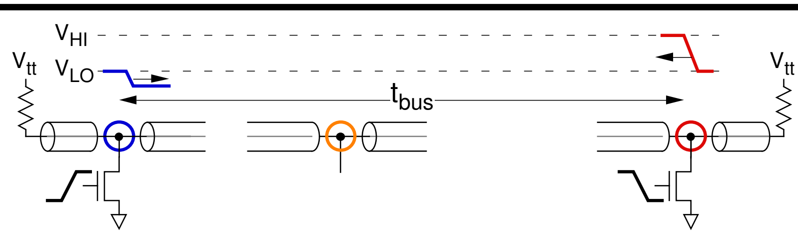
Driver has to sink twice the current as in a push-pull driver to get same signal voltage

Some example signaling systems:

Family	V_H	V_L	V_{ref}	assumed Z_{bus}	I_{OL}	(two term's) P_d(term)	P_d(tx)	Remarks
BTL (Backplane Transmission Logic)	2.1	1.0	1.55	30Ω	73mA	80mW	73mW	Bipolar current-mode
GTL (Gunning Transmission Logic)	1.2	0.4	0.8	30Ω	53mA	42mW	21mW	Voltage Mode
RSL (Direct RAMBUS)	2.5 1.8	1.5 1.0	2.0 1.4	50Ω 50Ω	40mA 32mA	40mW 26mW	80mW 32mW	Current Mode

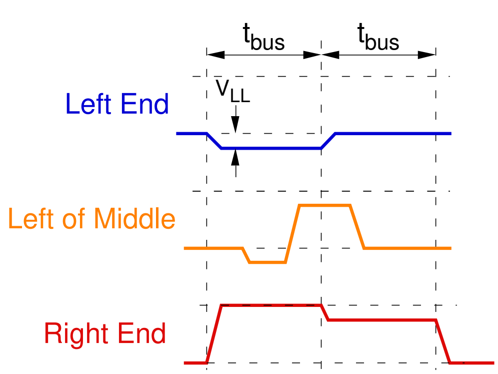
Note: power is dissipated only when driving LO, so, averaged over 1's and 0's, actual operating power is about one-half of these values.

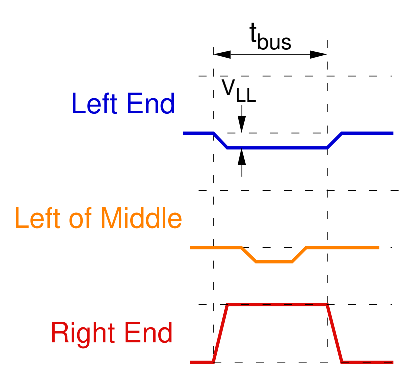
Wired-OR Glitch



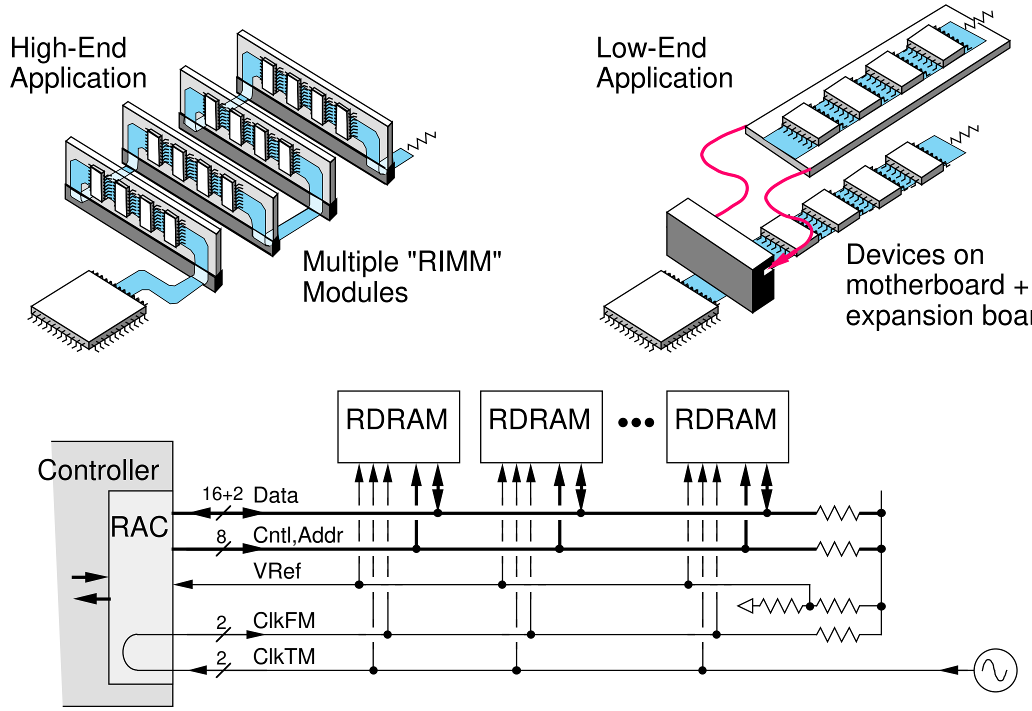
Wired-OR scheme with "voltage-mode" drivers requires two full bus transit times to settle the 'glitch'.

Current-mode drivers get it done in one bus transit time, because no reflections off the drivers.

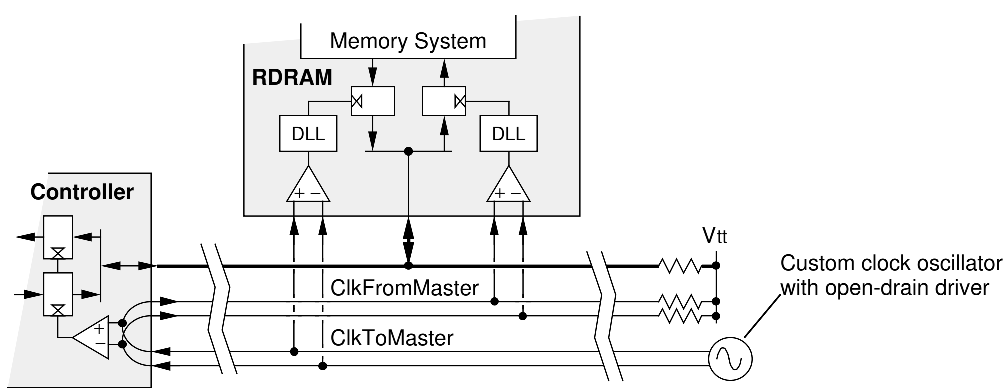




Direct RAMBUS (RDRAM)



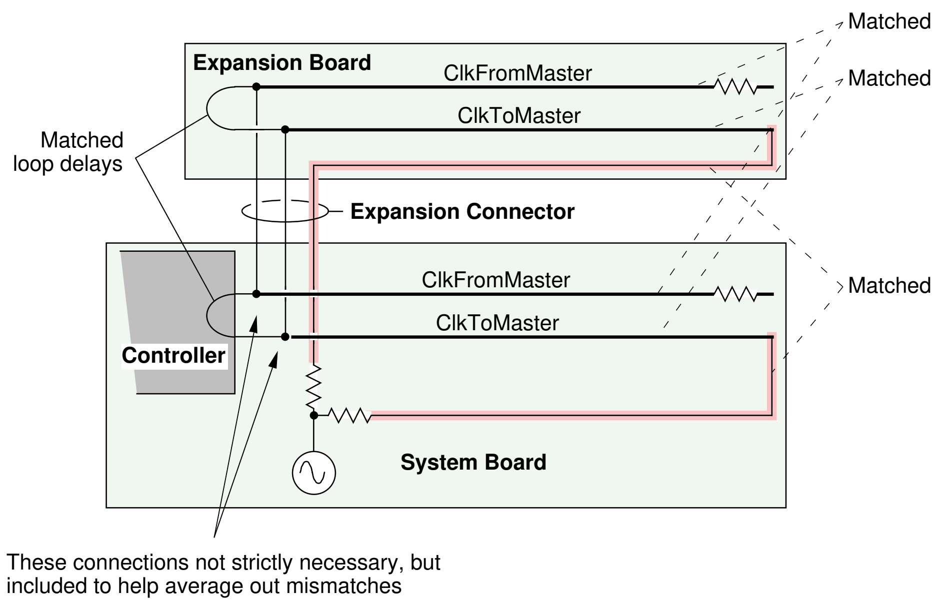
RDRAM Clocking



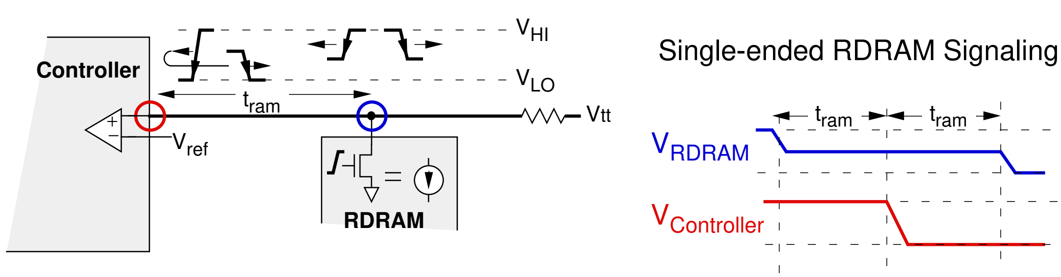
A variation on source-synchronous timing:

Clocks travel in same direction as data for both types of transfers

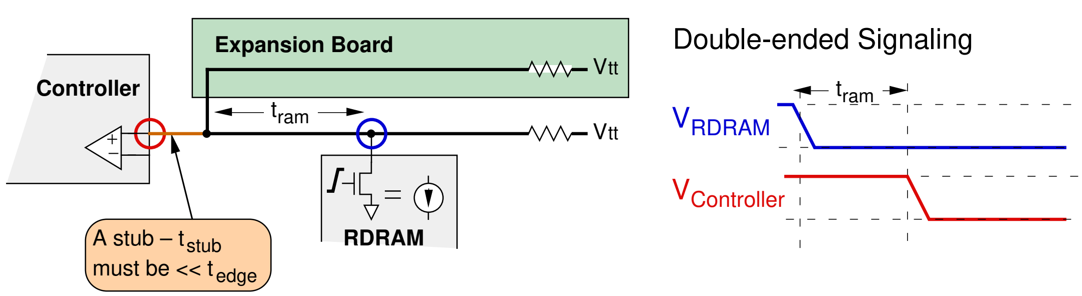
RDRAM Clocking with Expansion Board



RDRAM Signaling



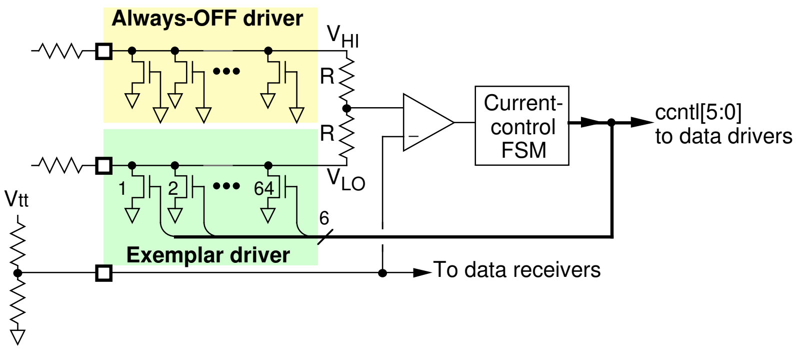
A variation on source termination; works because RDRAM slaves only talk to Controller



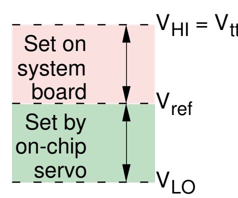
No half-step, because I_{OL} automatically cranked up to take care of two terminators

RDRAM Current Control

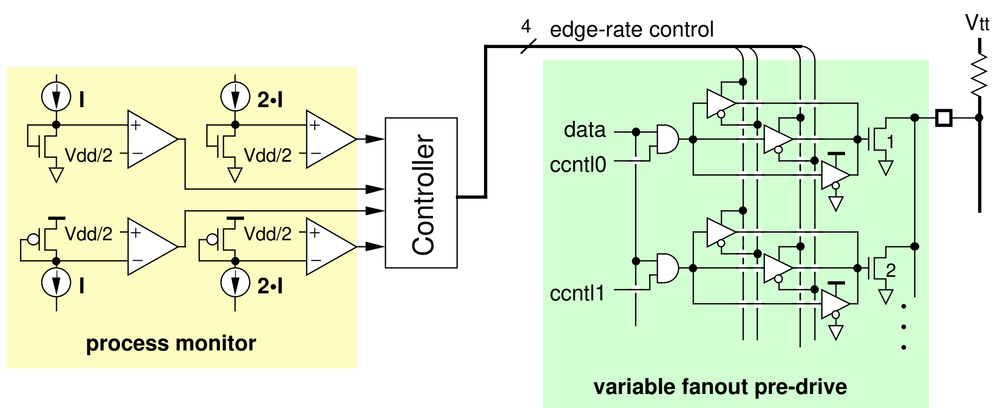
Drive-Current Control Servo



Current can be adjusted in ± 0.5 mA steps

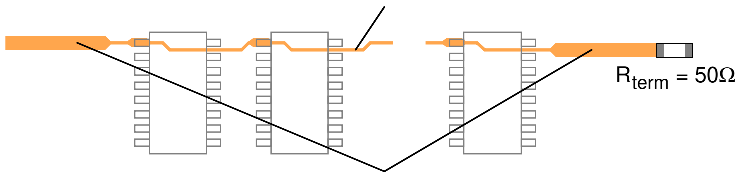


RDRAM Edge-Rate Control



RDRAM (and others) PC Board Layout Concerns

Loaded sections of line drawn so that
Z_0(bare) is high (~95Ω) allowing Z_{bus} ~ 50Ω



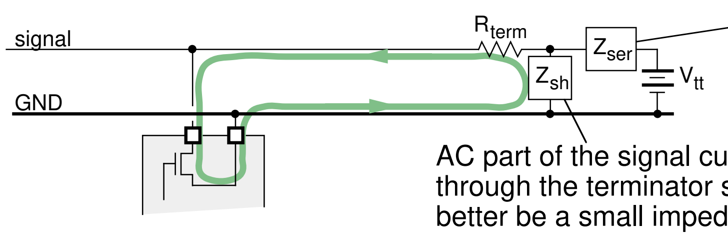
Unloaded sections drawn so that Z_0(bare) ~ 50Ω

Can use surface traces (microstrip) to get high bare impedance, BUT...

...there may be **lots** of forward crosstalk!

- Forward x-talk gets worse with line length ('piles up')
- ..and gets worse with faster edge rate.
- High-impedance means low **C**, so possibly lots of backward x-talk, too!

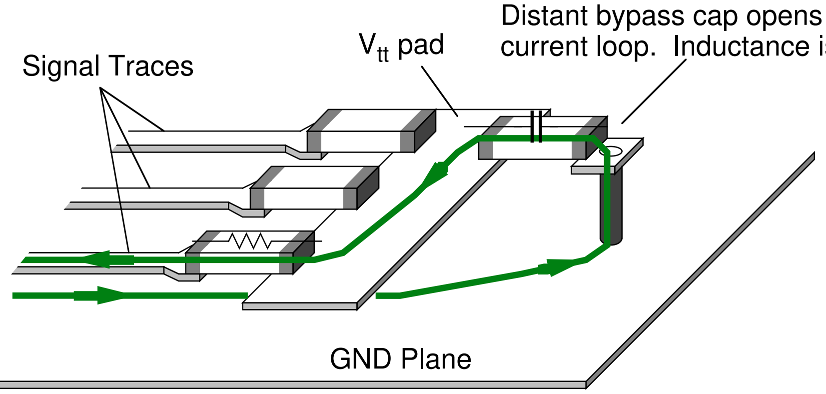
Terminator Layout



Terminator supply may be distance, so large series impedance

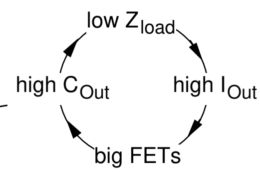
AC part of the signal current flows through the terminator shunt, so it better be a small impedance!

Distant bypass cap opens up large current loop. Inductance is bad!



Summary

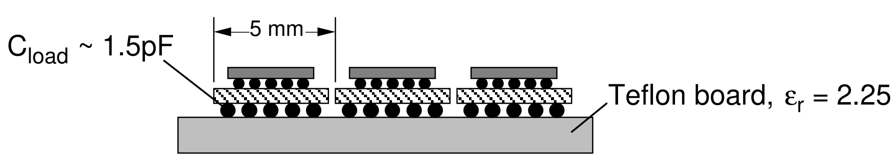
- A bus is a T-line with periodic discontinuities:
  - Lower impedance
  - Lower wave velocity
  - Lower bandwidth, cutoff frequency
- Bus transmitter has to drive $Z_{bus}/2$
  - the "spiral of badness"



- To break the spiral, need lower signal voltage swing
  - improved reference and receiver
  - smaller driver, lower drop capacitance
- Two techniques currently in vogue:
  - series stub termination (actually *isolation*)
    – a surprisingly effective technique
  - "open-drain" drivers, terminated to low V_{tt}
    – driver must supply 2x current, but only consumes power driving LO
    – current-mode version avoids the dreaded wired-OR glitch

Future Prospects for Bus Designs

Cutoff frequency is a fundamental limit, not power, so want low impedance, small t_{fs}



$$B_f = \sqrt{1+C_{Load}/\mathbf{C}\Delta x} = \sqrt{1+C_{Load}/C_{Seg}} \sim 2$$

$$F_0 = \frac{1}{\pi t_{fs} B_f} = \frac{1}{\pi \cdot 25\text{psec} \cdot 2} = \underline{6.4\text{ GHz}}$$

$$Z_{Bus} = \frac{50\Omega}{2} = 25\Omega$$

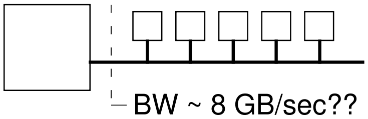
So, even pulling out all the stops, it's hard to see how F_{bus} could get much higher than about 1 GHz (2 Gbits/sec/wire).

Could reduce power by reducing voltage swing.

Differential current-mode bus drivers could go down to perhaps 50 mV swing (100 mV signal). Gets rid of troublesome V_{ref}. Perhaps 10 mW per driver, but costs two wires per signal.

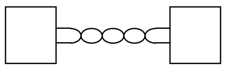
Busses? Just Say No.

Busses:



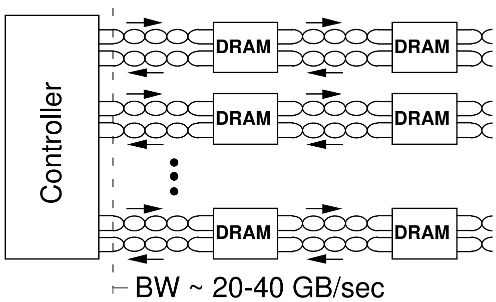
Today: 30 wires @ 800 Mbits/sec, ~3 GB/sec
Tomorrow: 30 wires @ 2 Gbits/sec, ~8 GB/sec ?

Serial Links:



Today: 1-2 Gbits/sec/wire-pair fairly common
4 Gbits/sec/pair demonstrated
Tomorrow: 10-20 Gbits/pair clearly feasible

Future Memory Interconnect??



Take those 30 wires and build 8 bidirectional serial links.

Bandwidth scales with technology, until fundamental limits reached.

Extensibility by simple linear network, but have to solve latency problem. Trees and other networks clearly feasible!

References

Textbooks & Paper Collections on Computer Interconnections

Bakoglu, H., *Circuits, Interconnections, and Packaging for VLSI* , Addison-Wesley, 1990, ISBN 0 201 06008 6

Dally, W. and J. Poulton, *Digital Systems Engineering* , Cambridge University Press, 1998, ISBN 0 521 59292 5

Johnson, H. and M. Graham, *High-Speed Digital Design–A Handbook of Black Magic* , Prentice-Hall, 1993, ISBN 0 13 395724 1

Lee, T. , The Design of CMOS Radio-Frequency Integrated Circuits, Cambridge University Press, 1998, ISBN 0-521-63061-4 (HB) 0-521-63922-0 (PB)

Tewksbury, S. (Editor), *Microelectronic System Interco nnections–Performance and Modeling*, IEEE Press, 1993, ISBN 0 7803 0405 5

Classic Textbooks on Transmission Lines

Matick, R., *Transmission Lines for Digital and Communication Networks* , IEEE Press, 1969, 1995, ISBN 0 7803 1121 3

Dworsky, L., *Modern Transmission Line Theory and Applications* , Krieger Publishing Company, Malabar, FL, 1988, ISBN 0 89464 276 6

A Few Papers on Recent Advances in Memory Signaling

See last year's ISSCC Proceedings, Session 10!

Lee, B., et.al., "A 2.6-GByte/s Multipurpose Chip-to-Chip Interface," IEEE JSSC 33(11), pp 1617-1626

Griffin., et.al., "A Process-Independent, 800-MB/s, DRAM Byte-Wide Interface Featuring Command Interleaving and Concurrent Memory Operation," IEEE JSSC 33(11), pp 117411751

References

Web Resources

PC-100 & Related – http://developer.intel.com/design/chipsets/memory/index.html#sdram

RAMBUS –
http://www.rambus.com
http://www.rambus.com/html/technical_papers.html

SLDRAM – http://www.SLDRAM.com/FAQ/fqindex.html

The SSTL spec (and other JEDEC specs) – http://www.jedec.org/

An example SLDRAM Device:

http://www.micron.com/mti/msp/html/sldramprod.html

Some example RAMBUS Devices:

http://www.hea.com/hean2/rambus/home.htm

http://www.toshiba.com/taec

This talk, and a more complete list of references at:

http://www.cs.unc.edu/~jp/isscc.html