3. Signal Integrity

1

- Analogue Issues: sensitivities and coupling mechanisms.
- Digital Issues: noise margins, supply decoupling, transmission line delays, crosstalk
- Packaging: supply bounce.
- PCB *vs.* IC issues
- Clock skew and clock circuit design.
- Thermal issues.

References

- P Horowitz & W Hill, “The Art of Electronics”, CUP 2nd edition 1989.
- J F Wakerly, “Digital Design”, Prentice Hall 4rd edition 2006.
- H Johnson & M Graham, “High-Speed Digital Design”, Prentice-Hall 1993.
- H Johnson & M Graham, “High-Speed High-Speed Signal Propagation”, Prentice-Hall 2003.
- R Morrison, “Grounding and Shielding Techniques”, Wiley 4th edition 1998
- J Buchanan, “Signal and Power Integrity in Digital Systems”, McGraw-Hill 1996.
- T Williams, “The Circuit Designer’s Companion”, Butterworth Heinemann, 1991.

2

Horowitz & Hill (Sections 7.23-7.25 & 8.33-8.35) is a very good starting point and Wakerly (Section 11.4) also gives a useful introduction to the digital stuff.

Johnson & Graham’s first book is my favourite for the rest, though you will have to be quite selective in your reading.

Morrison is excellent on the analogue side, but this forms rather a small part of this course (read it if you are doing any sensitive analogue circuitry).

Williams & Buchanan are also good, practical texts.

Analogue *vs.* Digital

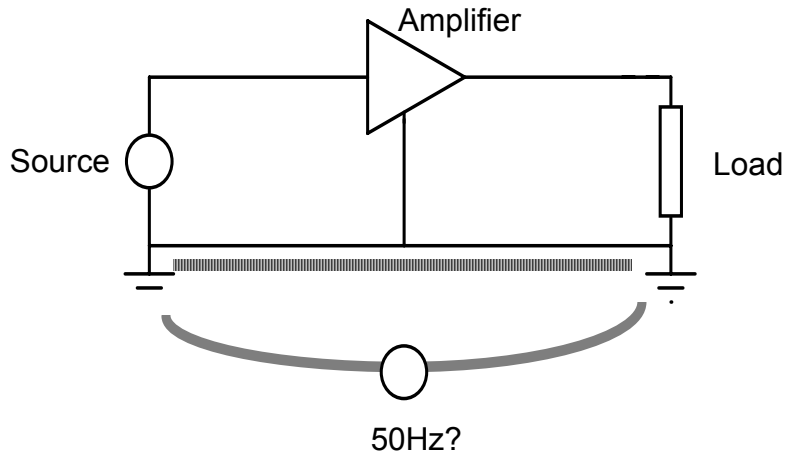
3

So far in this course, we have looked at how digital systems are made and at how they are designed allowing for the essential electrical issues such as the delay. However, whenever we actually build a system, we inevitably introduce extra components which are not on our circuit diagram. E.G. any two wires will have capacitance between them – any loop will have inductance. We need to be aware of the issues and satisfy ourselves whether or not their effects can be neglected.

Although this course is about digital systems, the world is analogue and the problems are essentially analogue in nature. You may have chosen to study digital logic because you didn't like electromagnetism but I'm afraid that increasingly you cannot ignore it. However, the basic issues are pretty straightforward and in the unlikely event that you need to invoke Maxwell's equations, you'd probably better find out about the available simulators.

When we build a primarily analogue or digital systems different things become more or less important.

Analogue: Earth Loops



4

The two “earths” can NEVER be exactly the same. If we measure the difference in the lab we will CERTAINLY see a component at 50Hz.

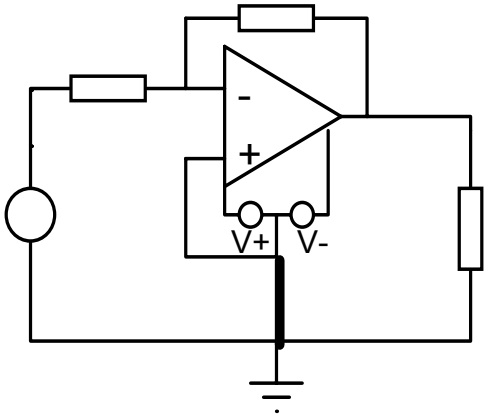
The resistance of the earth wire is critical here.

The problem arises because there is more than one earth connection.

Solution: Use just one earth or else a differential amplifier looking at the signal relative to its “ground”.

[Note that at higher frequencies (see later) , wires have impedance R + $j\omega L$ and the ωL will in the end be more important than the R.]

Analogue: Shared Resistance - 1



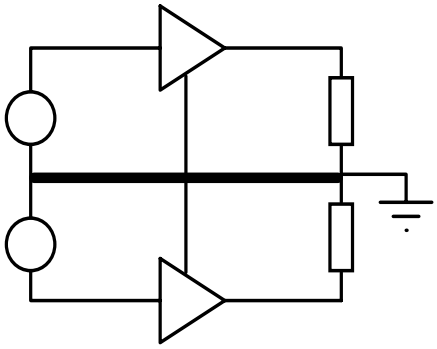
5

These circuits have similar problems. Although there is only one earth connection, two circuits share a common piece of wire.

Solution: Use separate wires

Use a “starred” earth connection

Analogue: Shared Resistance - 2



6

What is the problem here? How can we fix it?

.

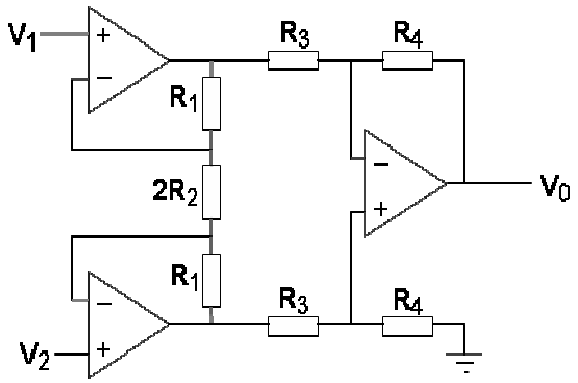
Conclusions

- Only ONE earth
- “Starred” connections to it
- Differential amplifier

7

Make your own notes.

An Instrumentation Amplifier



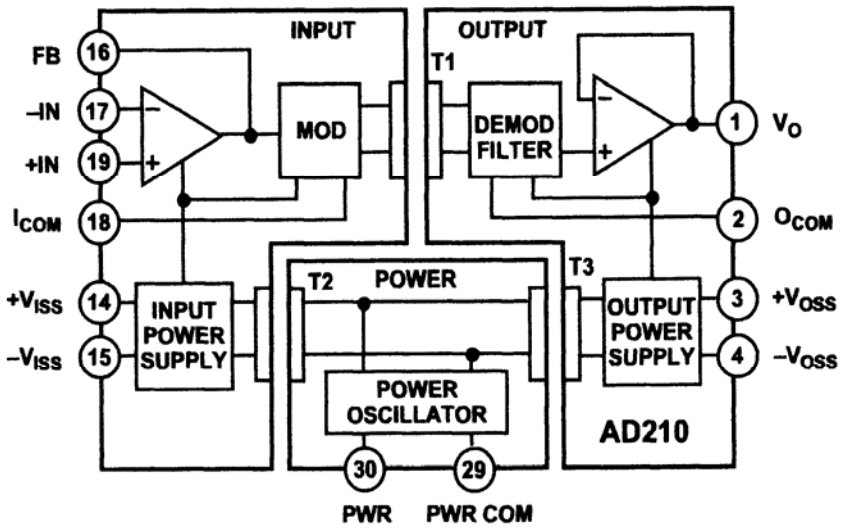
$$V_0 = \frac{R_4}{R_3}(1+\frac{R_2}{R_1})(V_2 - V_1)$$

8

By considering the input earth as another signal line, we can get around the problems of multiple earths.

This particular circuit is nice for analogue signals where we want a very high input impedance. Other circuits are better when the common mode signal may exceed the power supply rails.

Transformer-coupled Isolation Amplifier



9

By using an isolation amplifier we can remove ALL paths between the two earths.

This is useful where there is a VERY HIGH common mode signal (e.g. lightning strike in an aircraft).

The diagram shows the Analog Devices AD210 which is transformer coupled and also provides isolated power supplies. It is rated at 3500Volts and 20kHz. Other devices are available which use capacitive coupling and optical coupling.

Analogue: Capacitative Coupling

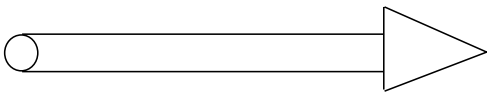
10

Sketch a picture of a piece of electronics kit (e.g. an oscilloscope) on the bench in the lab. There is “stray” capacitative coupling between all parts of our system.

The more sensitive the circuit and the stronger and nearer the source, the more of a problem it is.

However, we can shield the sensitive components with a “Faraday Cage”.

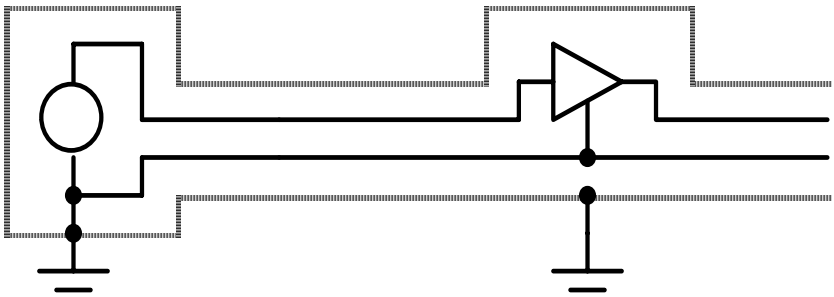
Analogue: Shielding



11

How do you shield this system?

Shielding - 1

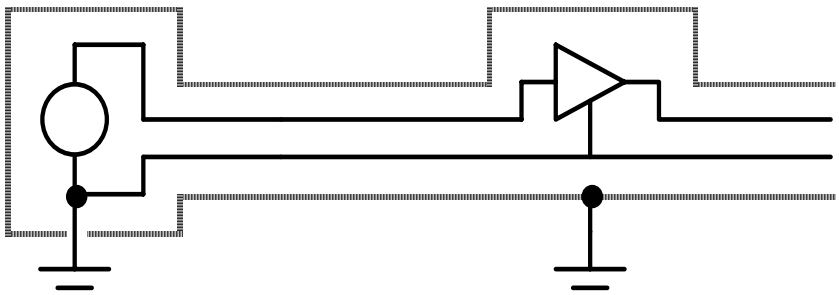


12

Identify the problem in this circuit and propose a solution.

Consider capacitance coupling between shield and circuit wires; current sourced by the difference in the two earths may flow through any closed path formed by wires and capacitance; if it flows along a signal wire we can get an IR voltage drop.

Shielding - 2



13

… identify the problem here and fix it.

Conclusions

- Conducting box
- Shielded wires / coax
- One earth
- Differential amp

14

Write your own notes.

Analogue: Inductive Coupling

15

As an easy illustration, most systems contain a mains transformer; at audio frequencies this may be the biggest source of noise (ever found/built a hi-fi amplifier which “hums”?). Make your own sketch.

Use a low leakage transformer (e.g. torroidal).

Keep it away from the most sensitive signals.

Keep instrument signal and its return wires close together and both paths short.

Use twisted wires or co-ax.

Magnetic shielding is possible but much more expensive than electrostatic screening. Low permeability materials are bulky and heavy and dealing with gaps more difficult.

Conclusions

- Control sources
- Distance
- Loop area
- Twisted wires / coax

16

Make your own notes!

RF Coupling

17

At high frequencies the “stray” inductance and capacitance are effectively increased (impedance jωL & admittance jωC). Even low resistance wires may have a significant inductance.

As critical parts of your circuit approach a significant fraction of a wavelength, they get more efficient as aerials.

The launching of an electromagnetic wave effectively involves both inductive and capacitative coupling. For the above solutions to be effective, the holes in the shield and the loop areas must be small fractions of the wavelength!

Digital Considerations

- Transmission Lines
- dv/dt, di/dt
- (Noise Margins)

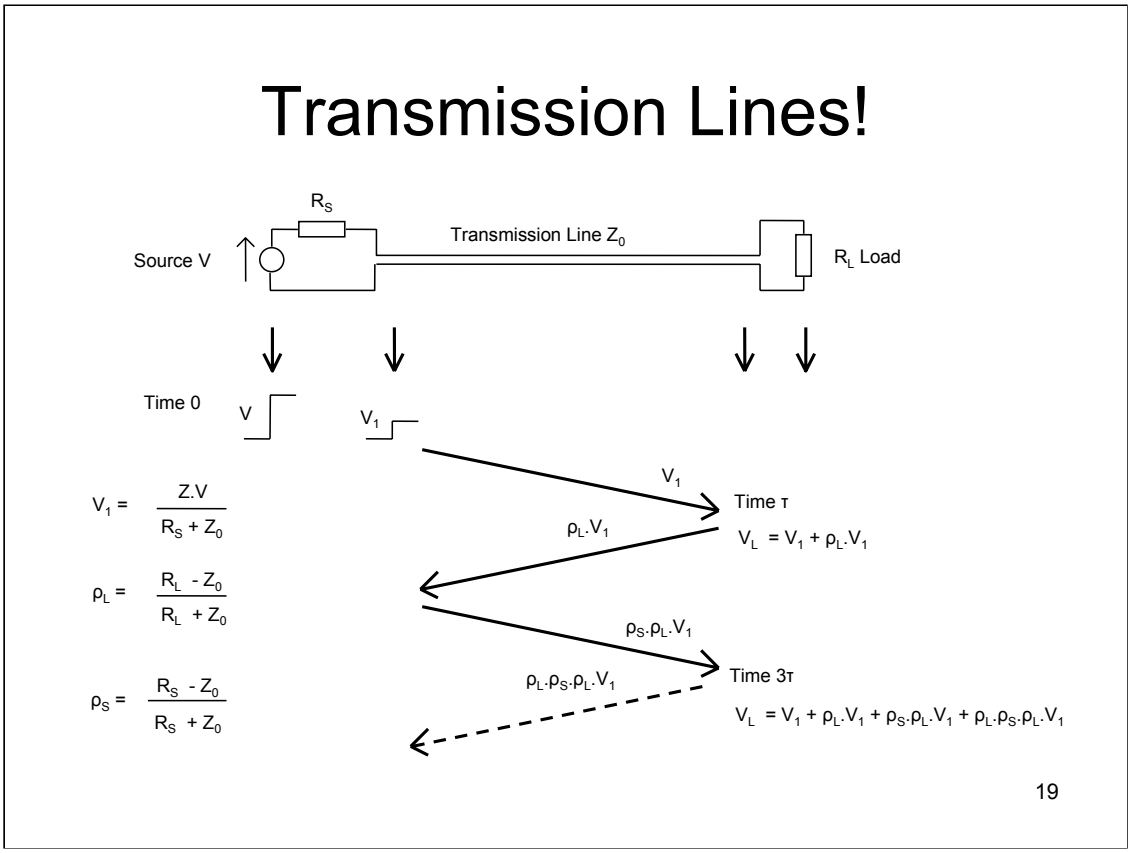
18

These days digital systems are fast and have many of the same problems as RF circuits. Just thinking about Fourier components of a square wave (say 3GHz), the LOWEST frequency component is the fundamental at 3GHz but we expect it to have some 3rd harmonics at 9GHz, 5th harmonic at 15GHz etc. For a pure square wave the nth harmonic is decreased in amplitude by 1/n relative to the fundamental. Triangular waves are better, decreasing at 1/n2 and intuitively we can expect our waveform to be somewhere in between these two cases (e.g. if we model it as a square wave with finite rise and fall times).

As a rule of thumb, we might consider the important bandwidth to be approximated by $1/(t_{rise} + t_{fall})$

However, the REDEEMING FEATURE of digital signals is our circuits have some NOISE MARGIN which avoids some of the problems associated with very small signals.

We will look at the above three aspects in turn.



Consider a digital signal applied to a transmission line.

The “characteristic impedance” of the transmission line Z_0, is a property of the physical layout of the line and establishes the relationship of voltage and current on the line in the short term before the input change has had time to travel to the far end and fine out what the load really wants.

We get reflections (just like light passing through glass), which depend on the relationship of Z_0 and the load impedance. $R_L > Z_0$ sends back a positive reflection cancelling some of the power in the original wave. $R_L < Z_0$ sends back a negative reflection asking for more.

Typical Velocities

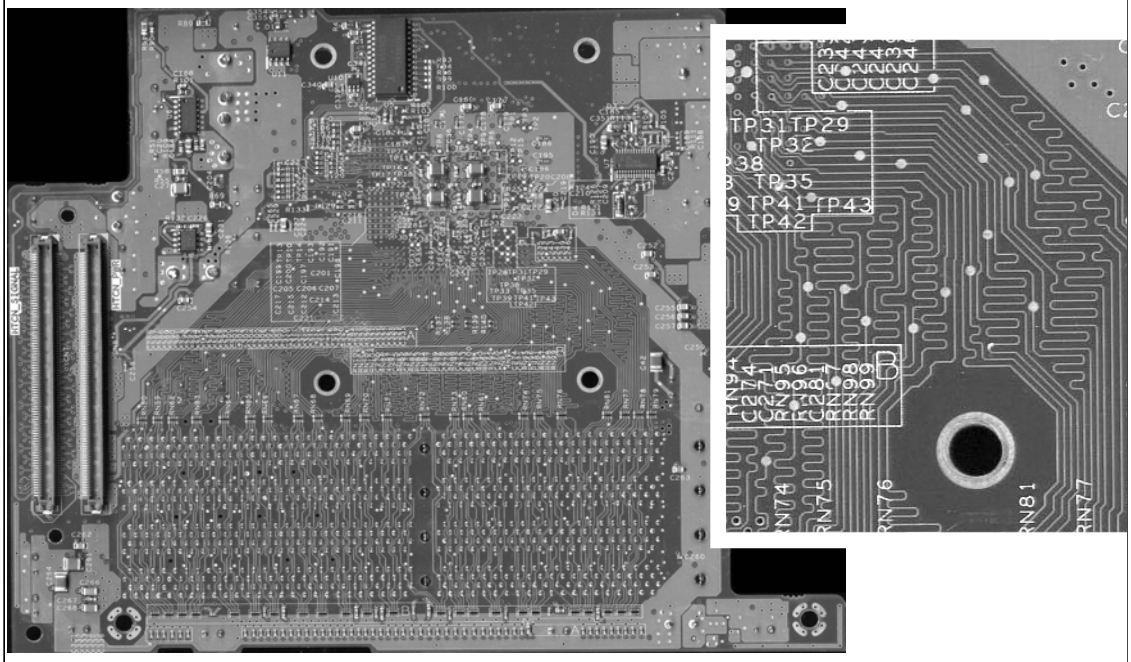
- Light: 30cm/ns
- Coax: ~ 20cm/ns
- PCB: ~ 15cm/ns
- Ceramic: ~ 10cm/ns

Depends on permittivity (& permeability).

20

$v = 1/\sqrt{(\varepsilon\mu)}$.

For modern digital systems, the propagation delay between and within PCBs has become a significant factor.



21

See on this motherboard how the delays to the memory are equalised by ensuring the tracks all have the similar lengths.

Typical Impedances

- Coax: ~ 50Ω
- Twisted pair: ~ 100Ω
- PCB track (et): ~ 40Ω

Depends on geometry and permittivity.

22

These are some typical characteristic impedances of our “wires”.

If the medium is “lossless” (low resistance wires and low leakage between them) the characteristic impedance is real (resistive):

$Z_0 = \sqrt{(L/C)}$ - see HLT p154 for different geometries.

The text by Johnson & Graham gives some others.

If also the source and load impedances are real and the input is a square edge, it is easy to predict the waveform:

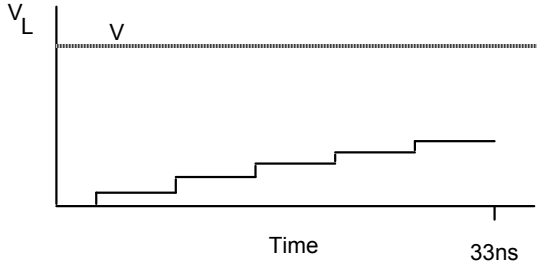
Example 1

Consider measuring a *high impedance* signal on a scope:

$Z_0 = 50\ \Omega$, $\tau = 3$ns, $Z_L = 1$M Ω, $Z_S = 1$ k Ω (say).

Hence $V_1 = 0.047$V, $\rho_L = 0.9999$, $\rho_S = 0.905$

and V_L = 0.09V, 0.18V, 0.26V, 0.33V, 0.39V, ... 1.00V



A 1MΩ scope input give almost perfect reflection to the 50Ω lead (the reflection coefficient is about 0.9999!)

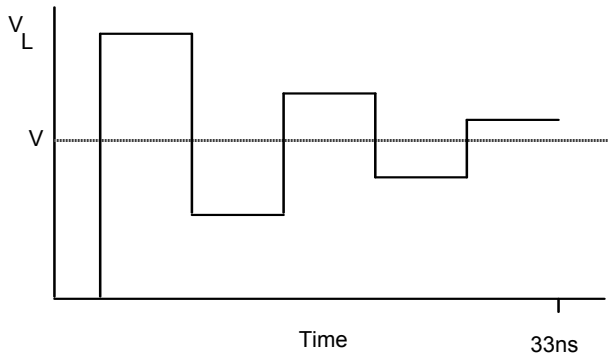
Consider here a typical scope lead with a 3ns delay time with a 1V square step and a (highish) source impedance of 1kΩ.

See how slow the edge becomes!

Actually in this case the 1/e rise “time constant” is about 57ns corresponding to a bandwidth of only around 17MHz - not much hope of seeing a fast edge here!!

Example 2

Consider measuring a *low impedance* signal on a scope:
$Z_0 = 50\ \Omega$, $\tau = 3$ns, $Z_L = 1$M Ω, $Z_S = 10\ \Omega$ (say).
Hence $V_1 = 0.833$V, $\rho_L = 0.9999$, $\rho_S = -0.667$,
and V_L = 1.67V, 0.55V, 1.29V, 0.80V, 1.13V, ... 1.00V



24

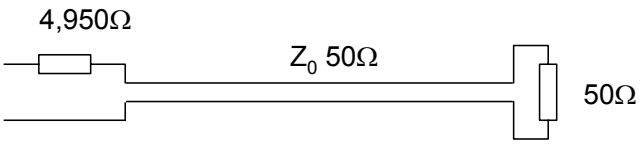
Here we see the effect of a lowish input impedance of 10Ω; more typical of the output impedance of a packaged IC. With only(!) a 5:1 mismatch, the time constant is about 15nsec (about 67MHz) but look at the waveform!

How many times have you looked casually at digital signals on a high frequency ‘scope and found the edges were “ringing”?

For other cases, try the Excel spreadsheet “reflections.xls” available at: http://www.weblearn.ox.ac.uk/bodington/site/mathsphys/engineering/cohort2001/4thyear/c3bvlsild/c3bvlsild_ln/.

Matching

- For high speed and a few cm or more, we must match one or (preferably) both ends
- Example: a x100 “50Ω divider probe”:



25

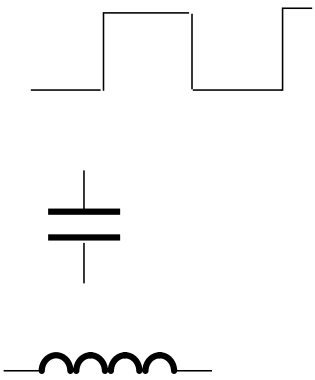
It is very easy to make up a probe like this.

Digital signals are generally big enough and strong enough that the 100-fold attenuation of this 5kΩ input impedance are not problems.

If you have the money, Tektronix (for example) will sell you a “x10” probe with 9GHz (AC Bandwidth) /500Ω and a “x100” probe with 3GHz /5kΩ.

Alternatively you spend rather more money and buy an “active probe” with an FET amplifier near the probe tip. The amplifier can have a high input impedance and its output is matched to the lead. Tektronix again sell 4GHz/100kΩ or 1GHz/10MΩ.

Frequency! di/dt! dv/dt!



26

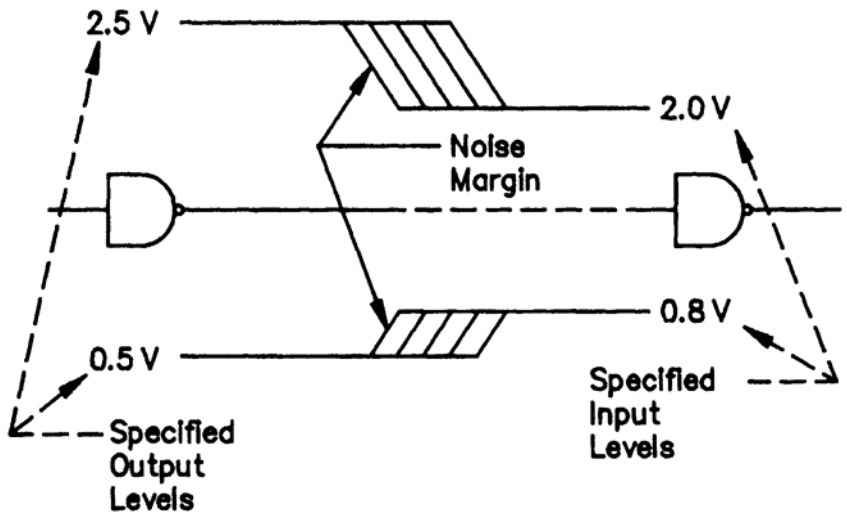
The other problems in digital systems are all caused by the rapid changes in voltage and current which hugely exaggerate the effect of any unwanted L & C.

Even a short length of parallel conductors will have some capacitance between them.

Even a small loop of wire will have some inductance.

How many harmonics do we need to represent a nice square wave?

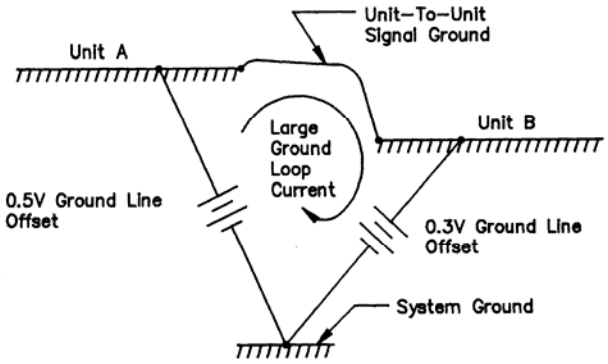
Noise Margins!



27

Noise Margins are often a few 100mV but they are falling as power supply voltages are reducing. Nevertheless, the existence of noise margins avoids the problems that we had designing with very sensitive analogue signals:

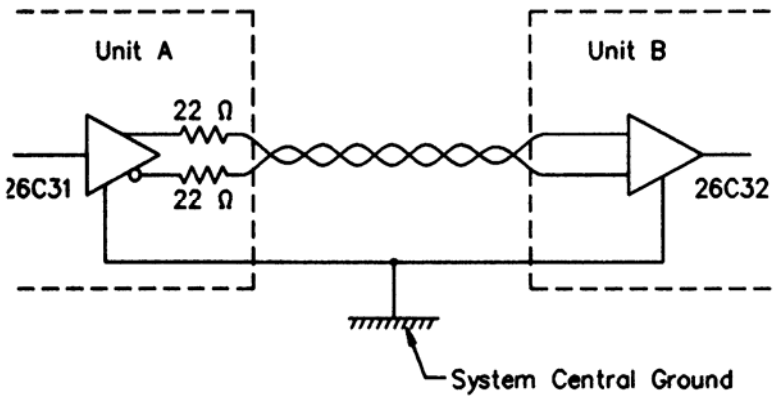
Digital: Ground Loops



28

Ground Loops have to be excited by pretty big signals before they become a problem in a digital circuit.

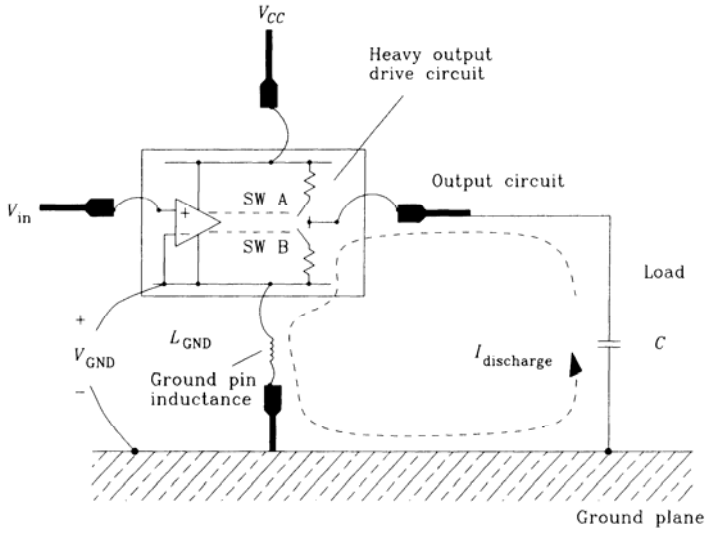
Line drivers/receivers



29

When taking a signal from one unit to another it is good practice to use a matched driver and a differential receiver.

Shared Inductance



$$V_{GND} = L_{GND} \frac{d}{dt}\left(I_{discharge}\right)$$

30

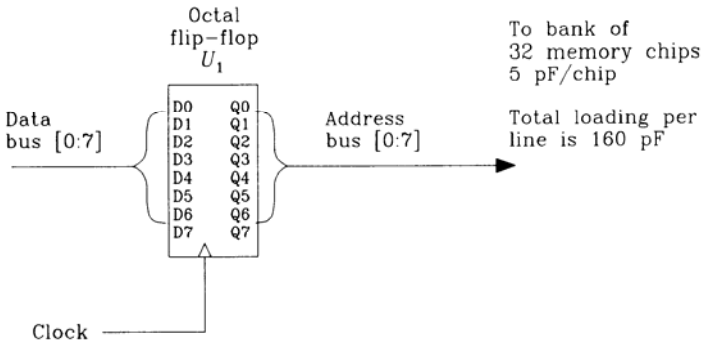
When sharing wires between digital circuits the inductance of the wires causes just the problems illustrated earlier for more sensitive analogue systems and shared resistance.

A voltage drop in a ground lead with inductance L is L.di/dt.

di/dt is biggest at the instant we start to pull down (pull up) and is given by V_{CC}/R. Making R pull-down (pull-up) bigger will reduce di/dt but will slow the edge down too (the discharge time has a time constant C.Rpull-down).

Inserting a series resistance will have the same effect. (Incidentally, this doesn't affect the overall power dissipation since the total energy thrown away per cycle is q = CV – half consumed when you charge C up & half thrown away when you discharge it.)

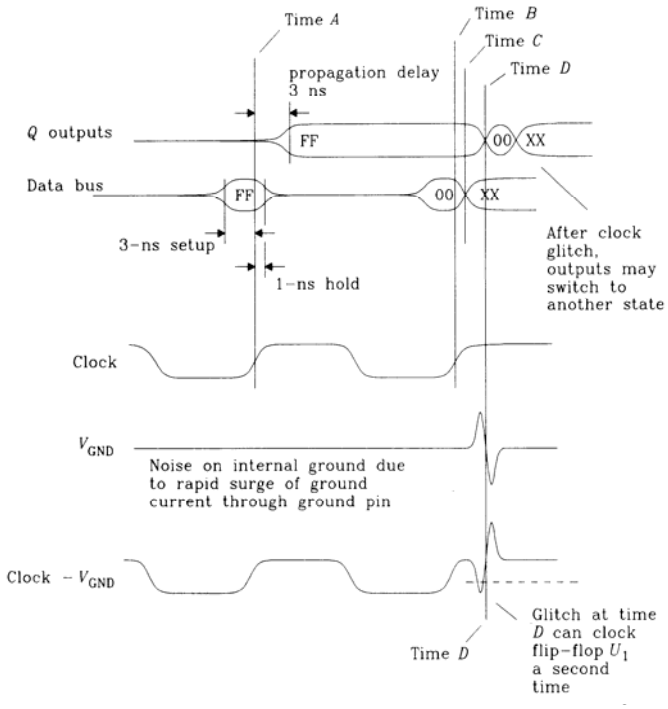
Ground Bounce - 1



31

Consider this register driving a heavy capacitive load.

Ground Bounce - 2



In this classic illustration of “Ground Bounce”, we assume that the input circuit “sees” the incoming voltage relative to its local ground which is disturbed by the di/dt of the outputs switching. Of course, depending on the technology, we may find that the input circuit is effectively referenced to power or ground (or both) and may therefore get problems of “Power Rail Bounce” too.

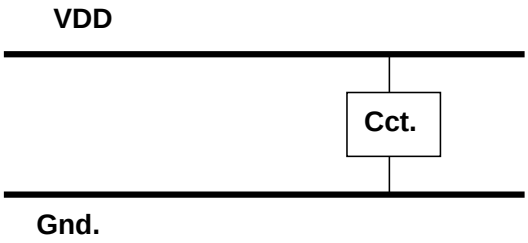
Limiting Ground Bounce

- Multiple pins
- Separate input & output grounds
- Low inductance packaging
- Use slower logic

33

Make your own notes.

Power Supplies



$$\int i.dt = q = C.v$$

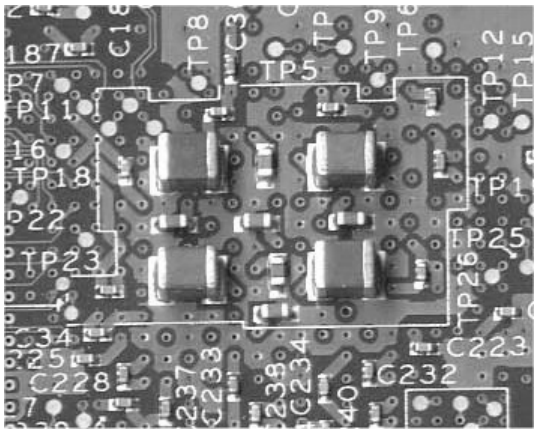
34

The big problem here is that the sudden current demands cause the supply voltage to droop which can affect the performance of ALL the ICs (thus causing more widespread problems than the ground bounce above).

Suppose the circuit is driving loads with a total capacitance of CL* at rise and fall times of T. If all the outputs change at the same time, within a time T we will need to supply a charge CL.VDD. Suppose T=1ns, and suppose our sudden demand can be propagated at 15cm/ns, the required charge must all be available within 15cm – WITHOUT the power supply drooping! I.E. If the power supply locally has a capacitance of CP, we need (say) CP>20.CL in order to get a droop of no more than 5%. That means the power supply must have a very low impedance and in practice, since there will be lots of other ICs around to worry about we must use “decoupling capacitors” on all ICs.

Typically 10nF is OK for small chips and 100nF for larger ones. More important than the value is that the capacitors should be suitable for high frequency operation (e.g. ceramics) and that they be very small with very small interconnections. (Any conducting loop will have inductance!)

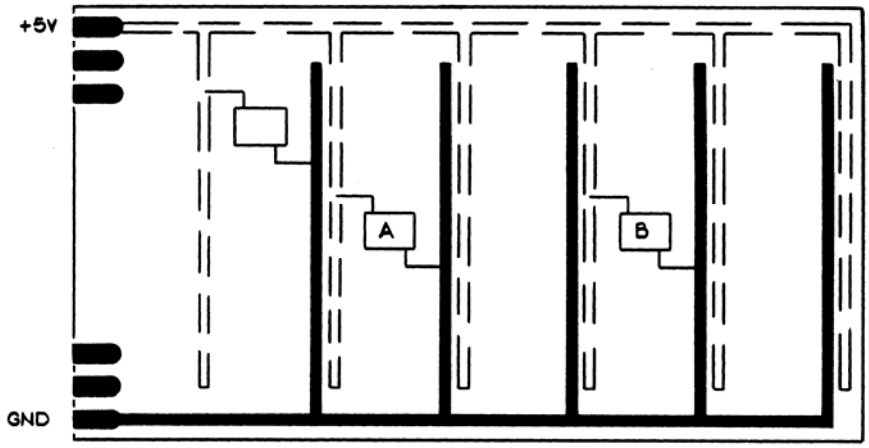
* Actually we will need to supply internal transitions too and some short-circuit currents during transitions. Hence it would be prudent at least to (say) double the value of CL in our calculations.



35

See here the board underneath an Opteron Processor?

Board Layout: Bad!

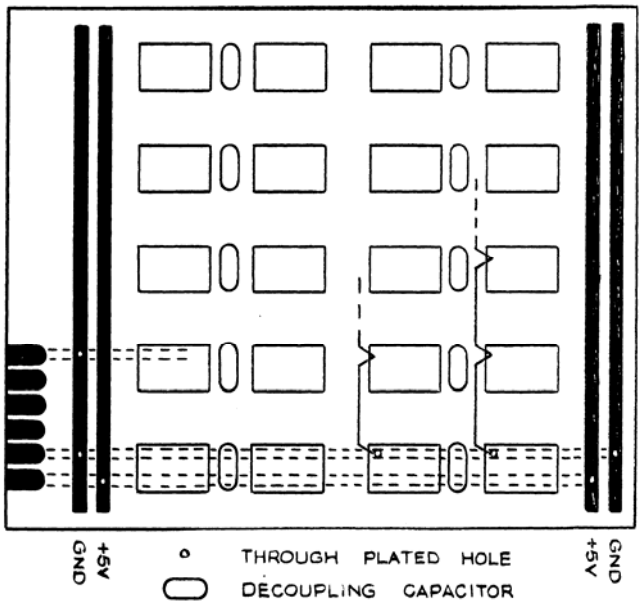


A BAD LAYOUT GIVING HIGH INDUCTANCE AND FEW ADJACENT SIGNAL RETURN PATHS, WHICH LEADS TO CROSS - TALK

36

Here there is a very large loop to supply current to the ICs and the characteristic impedance of the supply lines is as large as you can make it!

Board Layout: Better



37

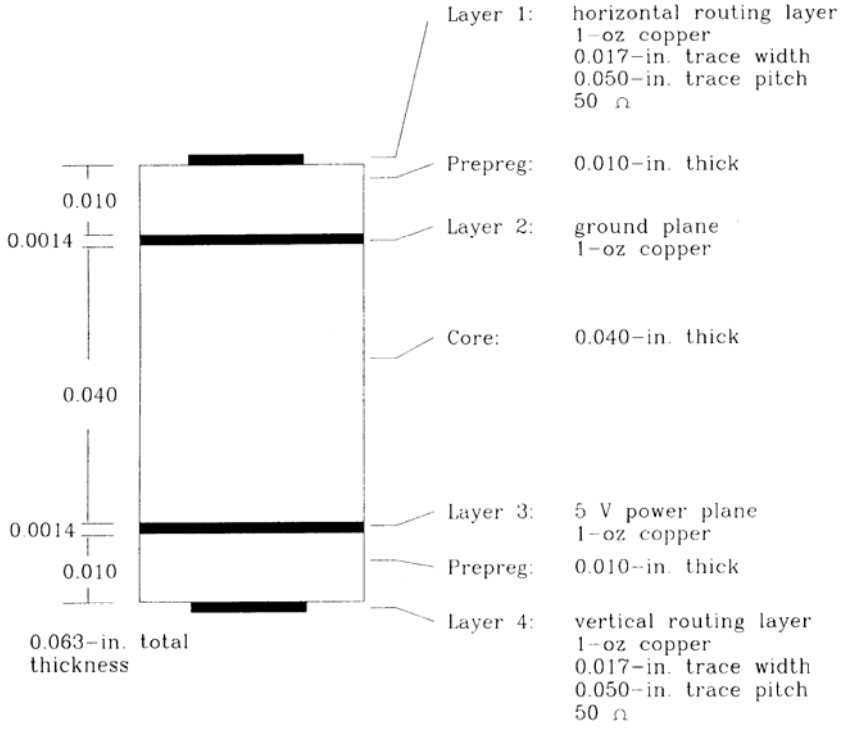
In digital systems it is VERY MUCH BETTER to have lots of connections to the supply rails and shortest possible distance between power and ground lines which lowers their characteristic impedance.

THIS IS IN COMPLETE CONTRADICTION OF THE GROUND LOOP/SINGLE EARTH PHILOSOPHY NEEDED FOR THE SENSITIVE ANALOGUE CIRUIT.

(And thus if you have a mixed signal system it is normally the case that the analogue ground and digital ground should be kept separate.)

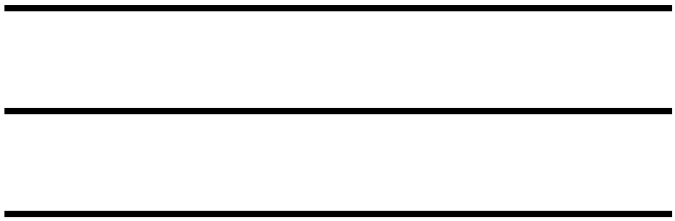
The lower the impedance the better. On top of that, add local decoupling capacitors as close as possible to every chip.

Best: Multi-laver Boards



Even better (and standard practice in industry) is to have entire power and ground planes.

Crosstalk

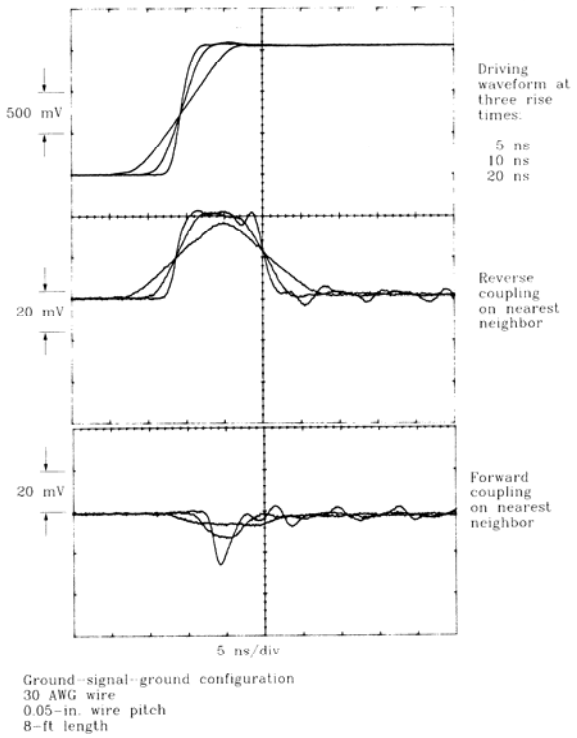


39

Cross talk arises from BOTH mutual capacitance and mutual inductance. The position of the signal return path is very important. E.g in ribbon cable, earth every other wire. On a PCB, run over a ground plane.

Estimation of cross talk is generally very difficult. Avoid running any sensitive, high impedance wires near any low impedance wires. Johnson & Graham give more guidance. An example is on the next slide:

Crosstalk in a Ribbon Cable



40

Here alternate wires in the ribbon cable are grounded. Note that the “reverse crosstalk” (i.e. with signals going in the opposite directions) is much worse than “forward crosstalk” so it’s best to give these even more separation.

dv/dt Control

- Slew rates

41

Since sudden voltage changes cause problems, it's best not to use faster logic transitions than you have to (but see also implications for low power later). Lots of FPGAs now have a programmable slew rate – use it!

Summary

- DC power rails
- capacitors have inductance
- signals take finite times
- and have reflections
- signals interfere with each other through
- resistive coupling
- capacitive coupling
- inductive coupling
- radiation

Frequency Scaling

- Everything gets worse as the frequencies get higher
- we need to worry about frequencies up to *at least* $f \approx 1/2T_r$
- That means for a 1GHz processor, where the rising edges are probably going to be around 0.1ns, we have to worry about frequencies of at least 5GHz.
- All this stuff gets ever more important as we delve further into “deep submicron” ICs

43

Make your own notes

Inside CMOS ICs

- Signal traces are very thin and have *appreciable resistance.*
- the inductance of signal traces is not usually important (yet – but see clock below) and the signal propagation is usually modelled by a "lossy" *R-C line*
- Circuit speed is now normally *dominated* by the R-C wiring delays

44

Make your own notes

CMOS ICs (cont.)

- for *very wide traces* (power and clock) the lines need to be modelled as R-L-C and transmission line reflections need to be considered.
- We need *in-circuit* power rail decoupling.
- As the geometry shrinks, the lateral dimensions get smaller but the vertical dimensions do not. We can have much more capacitance between traces than either have to the substrate. This is a recipe for *capacitance crosstalk.*

45

Make your own notes

Physical Issues: Conclusions

- Hidden components and connections
- In the extreme, digital is always analogue

- Problems are predictable
- Problems can be modelled

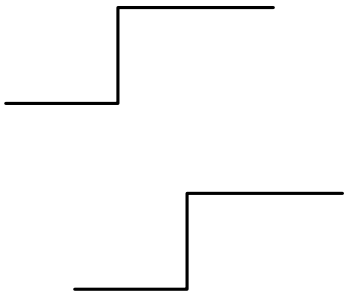
Clock Skew

- Clock rate? (t_{SU}, t_H, t_W, t_P)
- Clock skew!

47

In Section 1 of the course, we saw the importance of the set-up & hold times etc. of our flip-flops and how they influenced our design. Now that you have learnt that speed of modern digital circuits means signals may take a significant time to propagate around your circuit.

Timing Constraints -1



- Same clock edge at different circuits

  $t_{skew} < t_{Pmin} + t_{logic\text{-}min} - t_H$

- Premium on low t_H

48

Consider above the clock edge arriving at two different flip-flops with some clock skew between them.

Let's look the output of the top flip-flop and how it might affect the bottom one. This signal will be subject to a propagation delay in the top flip-flop and perhaps some further delay through combinational logic before it arrives at the bottom flip-flop. As discussed before, for pulse-mode (synchronous) logic to work, it is vital that the bottom flip-flop DOES NOT SEE THESE CHANGES until a hold time AFTER its clock edge.

This yields the constraint equation shown.

Timing Constraints - 2

- Consecutive clock edge at different circuits

  $t_{skew} < T_{CK} - t_{Pmax} - t_{logic\text{-}max} - t_{SU}$

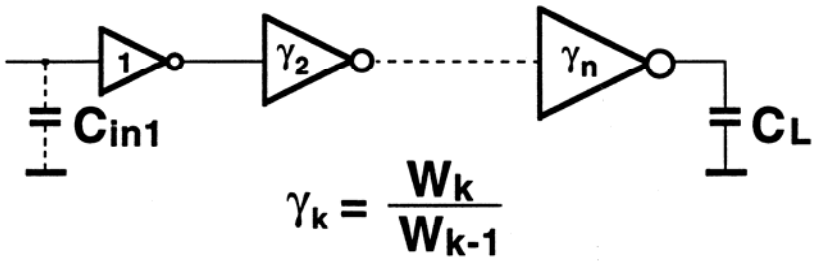
- Dictates Clock Rate
- Premium on low t_{SU}

49

Now consider the case of signals which we DO want to propagate. Suppose one flip-flop receives the top clock edge and another flip-flop receives the bottom clock edge AT THE NEXT clock interval. I.E. there is supposed to be a time TCK between the two edges. The output from the top flip-flop will be subjected to a flip-flop propagation delay and possibly further combinational logic delays but MUST ARRIVE at the bottom flip-flop a set-up time BEFORE its next clock edge.

This leads to the second constraint equation give above.

Clock Driver



$$\gamma_k = \frac{W_k}{W_{k-1}}$$

Uniformly tapering principle:

$$\text{If} \quad \gamma_2 = \gamma_3 = \gamma_4 = \ldots\ldots = \gamma_n = \gamma$$

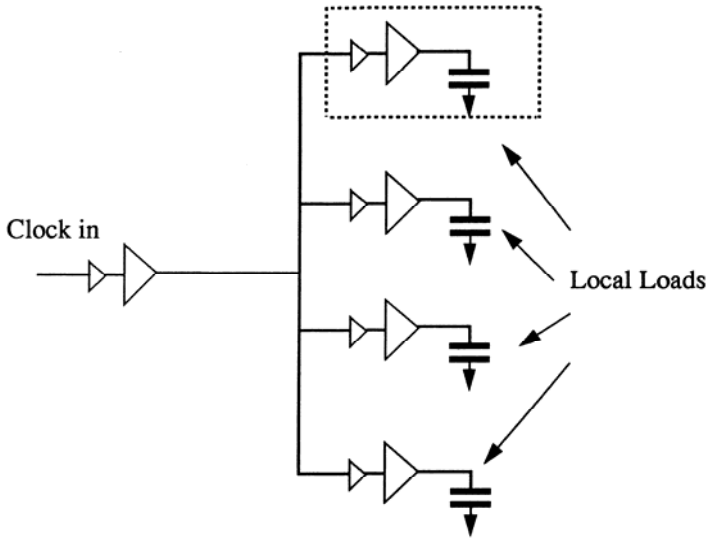
$$t_d = t_{dmin}$$

50

Given that the clock line goes everywhere, it typically has a high capacitance so we will need a low resistance driver. Typically we have to cascade cascade several inverters together as shown.

That's going to delay the clock quite a lot!

Clock Distribution

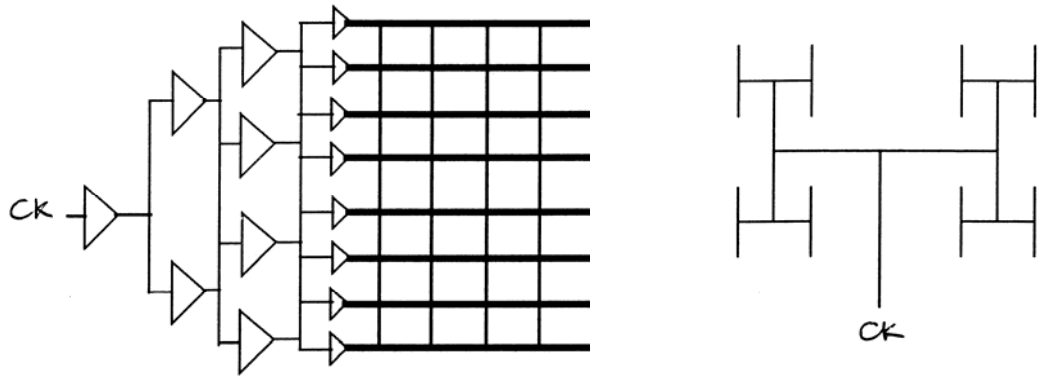


51

… and keep the clock trace wide to lower its resistance.

Here we distribute clock buffers through the design but this need great care to balance the branches if we are to avoid clock skew.

Clock Networks

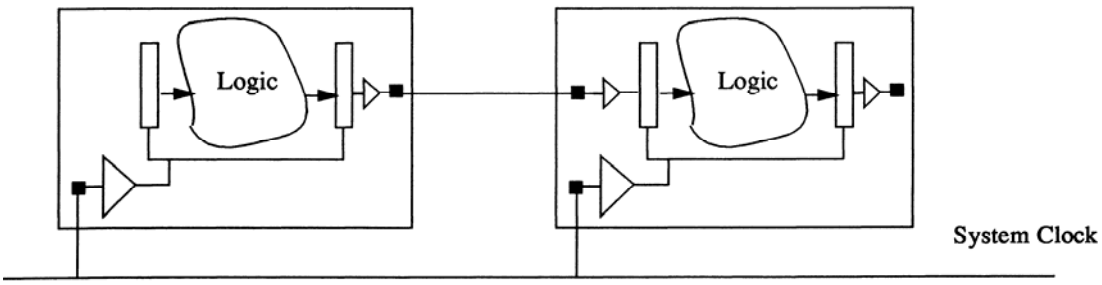


52

On the left is a “belt & braces” approach to avoiding clock skew and on the right the H-tree attempts to make all paths of equal length.

However we do it, we are going to delay the clock further – these techniques are simply going to reduce the local clock skew. That means if our signals are also local they ill be OK but if the go a long way around the chip we could be in trouble.

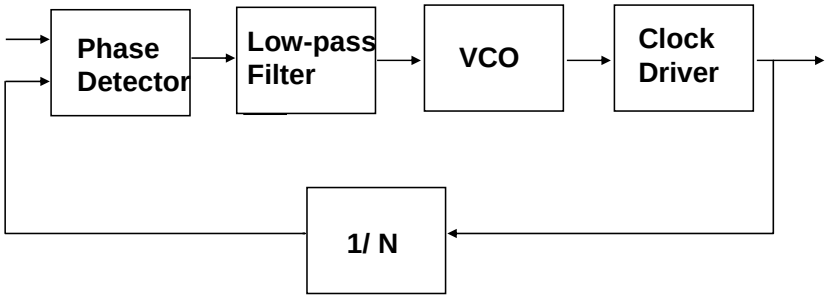
Chip-to-Chip Clock Delays



53

However well we do within a chip, the external connections and the delays of the different clock buffers in different chip designs is bound to introduce clock skew. We HAVE to do something about it!

Phase-locked Loop (PLL)



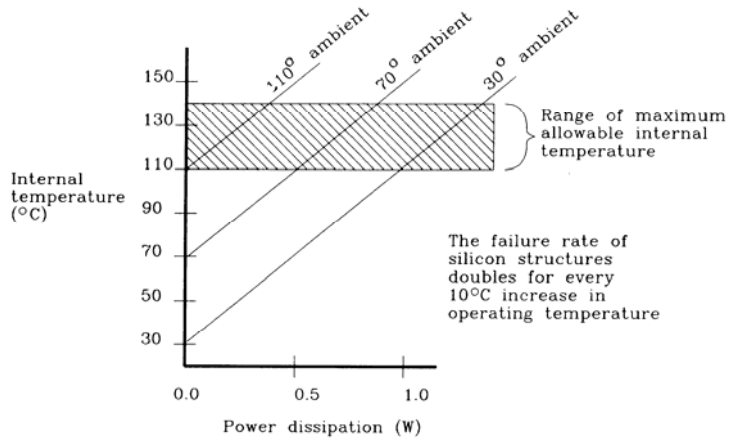
54

Whatever the clock delay in the clock driver circuitry (and the clock network), providing we know what it is, we can correct the timing with a PLL to give in effect a “zero delay”.

For this reason nearly all high-speed chips include a PPL in their clock circuits.

It is also commonplace to have several clock speeds in different regions of the circuit and to avoid distributing the fastest ones around the board. Thus if we have an input of 1GHz into the circuit above and N=2, we generate a 2GHz clock within the chip.

Thermal Resistance - 1



55

It turns out that the internal temperature of a chip is more or less a linear function of the chip-to-ambient temperature difference.

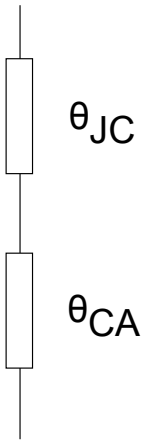
This leads to the concept of a “thermal resistance” in °C/W.

Of course in the end, high temperatures can damage the chip. Before this happens, the reliability os certainly degraded.

Thermal Resistance - 2

Θ_{JC} JUNCTION TO CASE THERMAL RESISTANCE	
16-pin plastic dual in-line package (DIP)	34°C/W
16-pin ceramic DIP	25°C/W
40-pin ceramic leaded chip carrier (LCC) with 10K square mil die	5.5°C/W
132-pin ceramic LCC with 50K square mil die	1.4°C/W

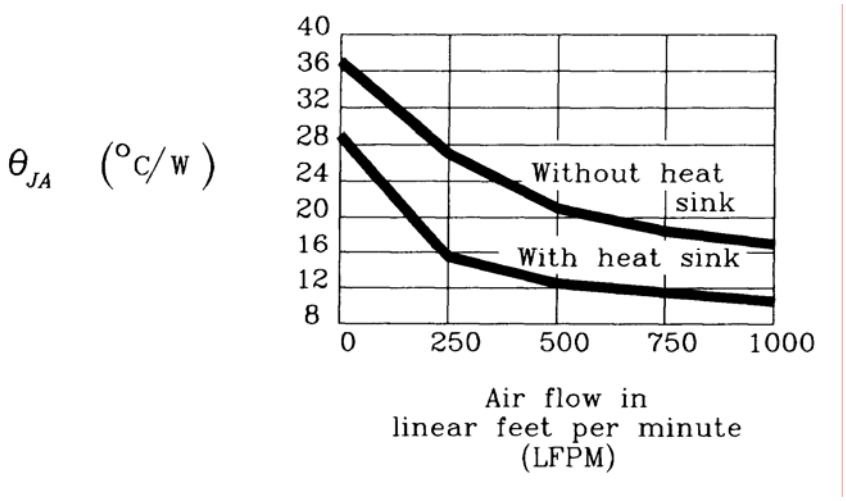
Θ_{CA} CASE TO AMBIENT THERMAL RESISTANCE	
16-pin dual in-line package (DIP) in still air	80°C/W
16-pin dual in-line package (DIP) in 400 ft/min air flow	35°C/W
72-pin ceramic pin grid array (PGA) in still air	34°C/W
72-pin ceramic pin grid array (PGA) in 400 ft/min air flow	18°C/W
72-pin ceramic pin grid array (PGA) in 400 ft/min air flow with heat sink	10°C/W



56

… and we can add up thermal resistances just like electrical resistance.

Thermal Resistance - 3



57

If needs be, we can add a heat sink to reduce θ_{CA} and/or add a fan.

In a student project you can calculate the thermal resistance you need and select a suitable heat sink.

In real life we might want to use a thermal simulator to model the heat sources and the air flow more accurately.