VLSI design

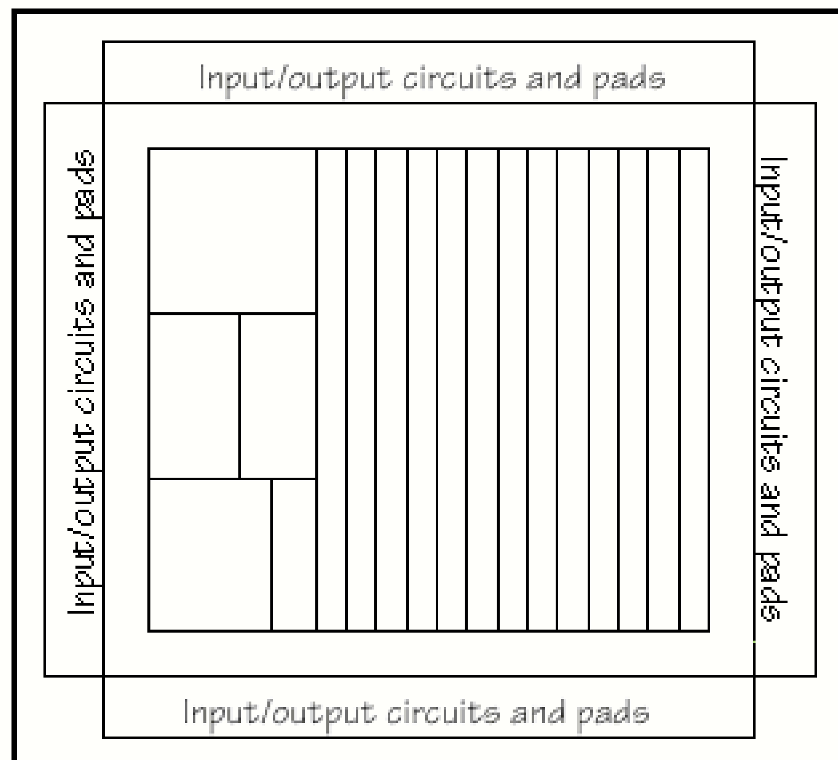
Floor planning and layout

P.J. Drongowski
SandSoftwareSound.net

Copyright © 1987-2013 Paul J. Drongowski

Plan ahead

- Good floor planning is like structured programming
  - Postpone tiny details until later
  - Design the overall system structure
  - Provide a context for cell (I/O) design
- Space and power are limited resources
  - Things never fit in the space allotted
  - Small size leads to good yield
  - Wiring occupies a large part of the real estate
- A good plan can minimize long, random wire runs
  - Manual routing is tedius and error-prone
  - Auto-routing cannot overcome poor placement
  - Length of time-critical signals must be controlled



Planning considerations

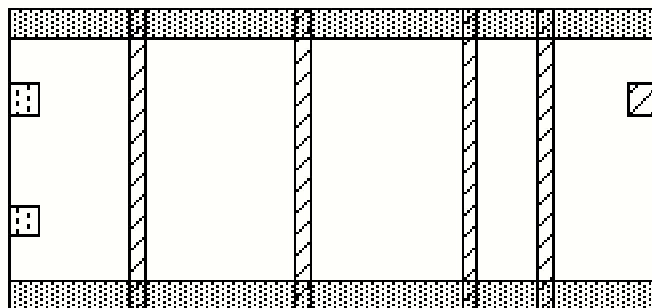
- Identify major functional blocks
  - Coarse grain functions, not gates
  - Minimize number of distinct cell / block types
  - Re-use cells if possible
- Estimate size of functional blocks
  - Use experience
  - Examine similar designs and systems
  - Prototype common or space-critical cells
- Place functional blocks within core
  - Avoid busses and random wiring
  - Use wiring by abutment
- External, off-chip connections
  - Short, direct connection to core logic
  - Pad assignment affects placement, vice versa
- Power grid
  - Must be routed on (sized) metal
  - Grid connects all subsystem
  - Minimize distances to keep losses low
- Clock distribution
  - Must (should be) routed on metal
  - Keep wire lengths short
  - Wire length affects clock skew and timing

Cell design

- Follow the chip plan
- Plan provides space and power budget information
- Cell electrical design
  - Determine maximum desired delay for cell
  - Estimate external wire length and loading
  - Choose transistor sizes and analyze (Spice)
    - Speed: Fast enough under estimated load?
    - Space: Too big for space budget?
    - Power: Is cell within power budget?
- May need to adjust chip plan
  - Shift needed resources from other subsystem cells
  - Reallocate saved resources if cell within budget

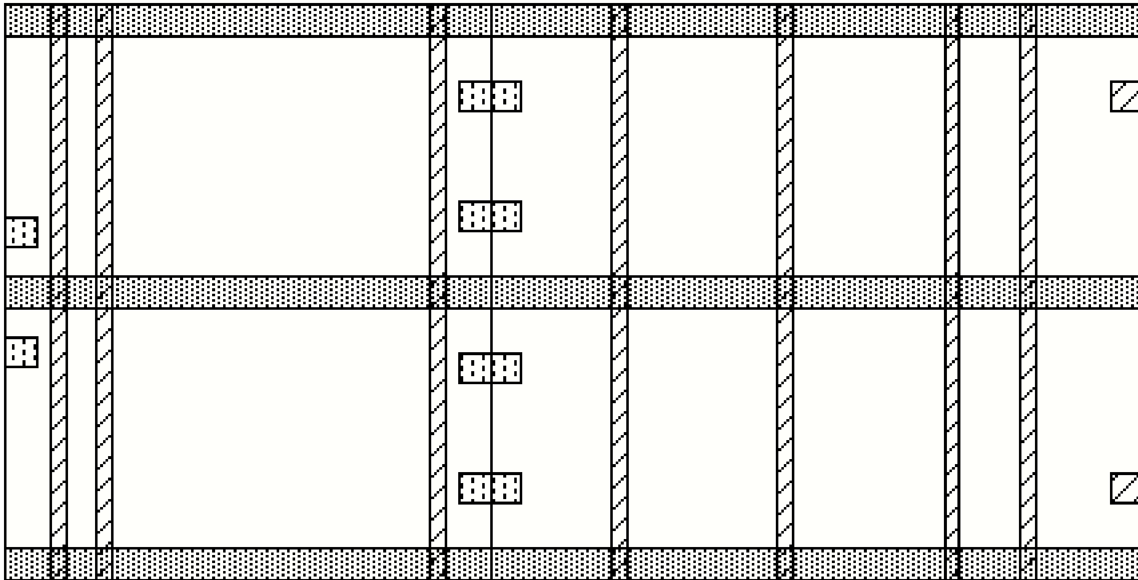
Cell layout

- Determine size, aspect ratio, pitch
- Identify inputs, outputs, power and ground
- Determine signal priority (metal, poly, diffusion)
- Hints
  - Data inputs at left, data outputs at right
  - Route control signals on polysilicon
  - Route power and ground on top / bottom edges
  - Control is perpendicular to data



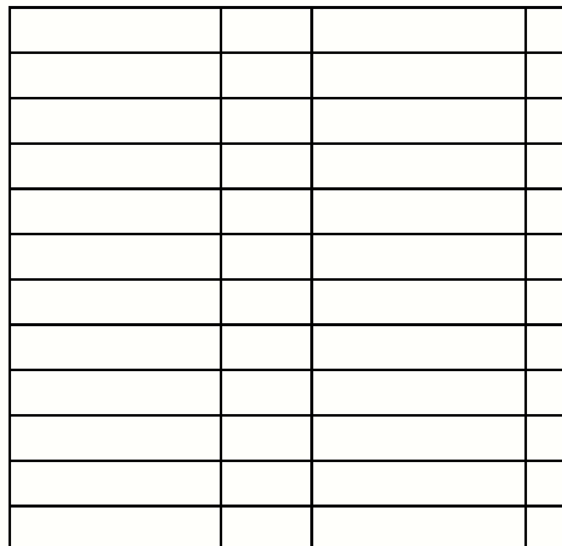
Abutment

- Match pitches (width)
- Align inputs, outputs and power busses
- MIrror to overlap power busses



Pitch and aspect ratio

- Assume a square subsystem shape
- Subsystem uses multiple cells, multiple bits
- Cells will be probably be long and thin



Four cells by 12 bits

Regularity

- Regularity begins with the choice of algorithm
- Example: Table look-up given a binary key
- Sequential implementation
  - RAM, address register, comparator, controller
  - Read RAM, compare, and increment address
  - At least four different cell types
  - Four separate blocks to be connected
- Content-addressable memory (CAM)
  - Large array of CAM cells
  - Each cell compares itself against pattern
  - One cell type replicated many times

Block placement

- Size of the individual block types
- Aspect ratio or shape of the block types
  - Square aspect ratio's for subsystems are best
  - Easier to pack square subsystems
  - Bit-sliced cells are often long and narrow
  - Necessary to get slices into square subsystem
- Opportunities for wiring by abutment
  - Abutment is always preferable
  - Minimizes both routing effort, length, area
- Length of interconnect to other cells
  - Keep wires short for high speed and low power
  - Critical path cells should be placed together
- Direct access to shared busses
  - Minimize the number of crossovers or unders
  - Use busses judiciously - they're slow and huge
  - Long parallel wire runs have crosstalk and noise
- Connection to I/O interfaces
  - I/O circuits and pads are always at chip boundary
  - Pads placed in fixed frame for auto-bonding
  - External pin-out is sometimes predetermined

Power grid

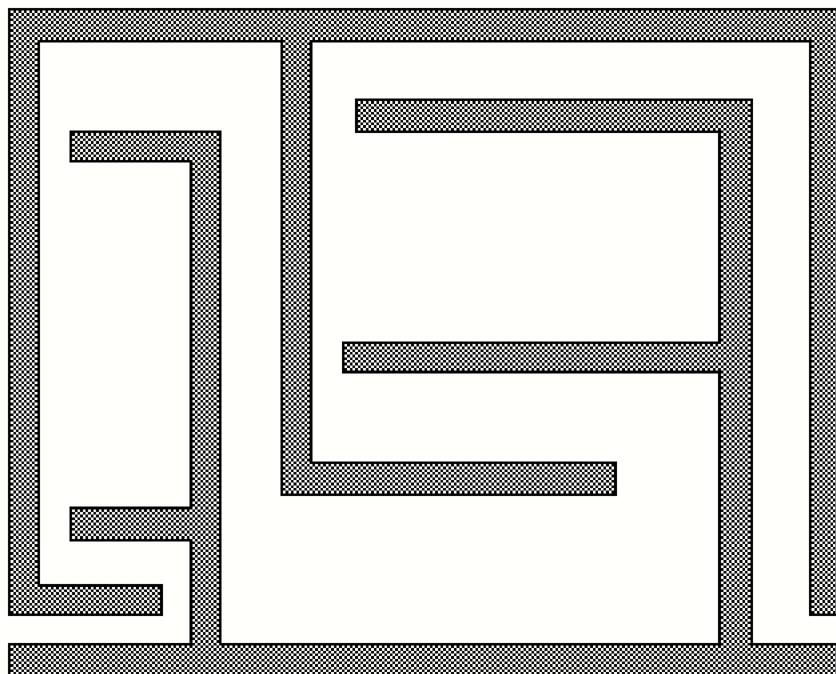
- Power and ground must be routed on metal
- No temporary distributions on poly or diffusion
- Paths enmesh to avoid crossing wires
- Metal wires must be properly sized

Power sizing

- Metal migration
  - Current density = J = A / square-micron
  - If J exceeds threshold, atoms physically move in direction of current flow
- Potential circuit failure
  - Atoms move faster at higher currents
  - Circuit will eventually blow like a fuse
  - Current density is highest at constrictions where cross section is smaller
- Aluminum wires
  - Maximum current density = 2 - 3 mA / square-micron
  - Use conservative limit of 1 mA / square-micron
- Capacity of typical minimum width wire
  - Feature size = 2 microns
  - Assume 1 micron wire depth, 3 lambda wire
  - Cross section = $2 \times 3 = 6$ square-microns
  - Maximum current = 6×1 mA/square-micron = 6 mA
- Estimating peak current
  - Both transistors are turned on
  - Resistive current path from Vdd to ground
  - Compute resistance through pull-up and pull-down transistors and apply Ohm's law
  - Current through minimum size CMOS inverter
    - Assume gate resistance of 10,000 ohms
    - I = V / R = 5V / 15K ohms = 0.3 mA

Power layout

- Route on metal over long distances
- Give ground priority over Vdd if necessary
- Minimize power voltage drop
  - Maximum voltage drop of 0.2 V (typical)
  - Voltage drop = $I \times R$
  - Avoid metal to metal contacts - high resistance
  - Try to stay on one layer (e.g., metal 2)
  - Keep power wires short
- n- and p-diffusion for power
  - Short, local connection to power rail
  - Use wide paths to lower resistance of connection
- Put output buffers and core logic on separate busses
- Tips for sizing the grid
  - Estimate current and size small subsystems first
  - Move up design hierarchy summing currents at each level
  - Size the power lines at each level

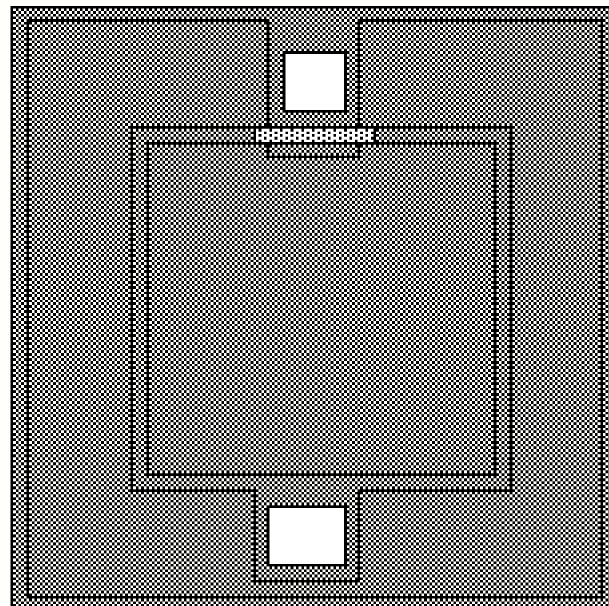


External off-chip connections

- Design I/O circuits to constant width and height
- Place pads to predetermined locations in frame
- Size and placement for automated wire bonding
- Auto bonding requires pad placement at edges
- Typical pad width is 100 μm to 150 μm
- I/O pads use common power busses
- I/O power is separate from logic power grid
- Vdd and ground form two rings around chip
- Use multiple Vdd / ground pads to reduce noise

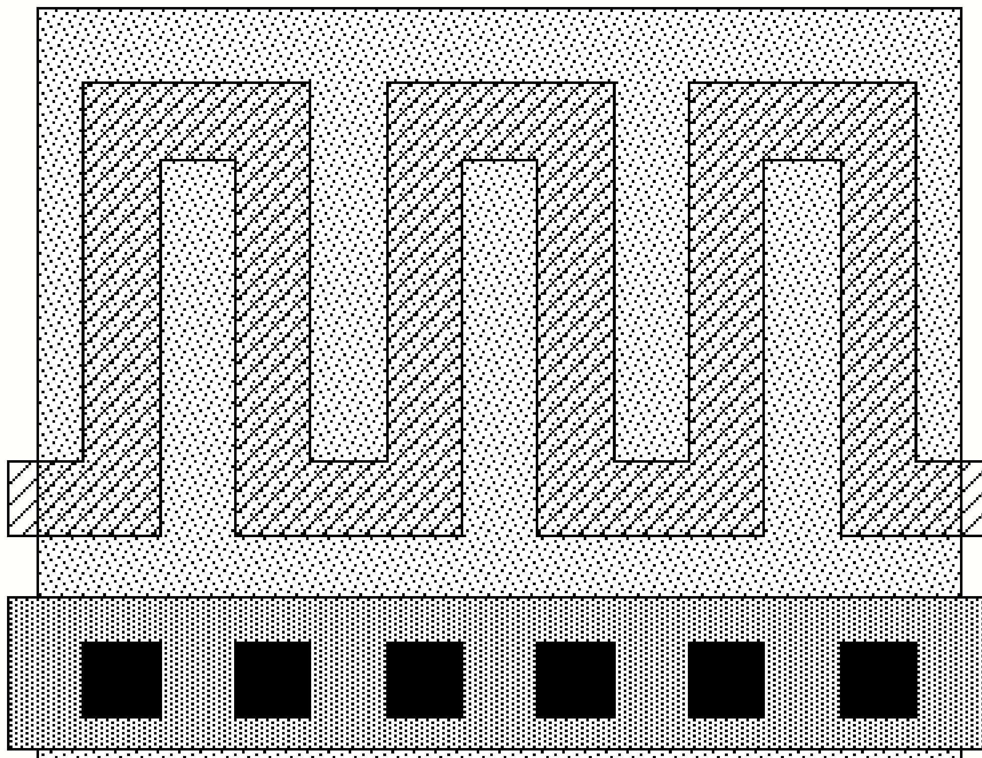
Power connections

- Power pad is a simple metal pad
- Requires opening in the overglass layer
- One rail must cross over



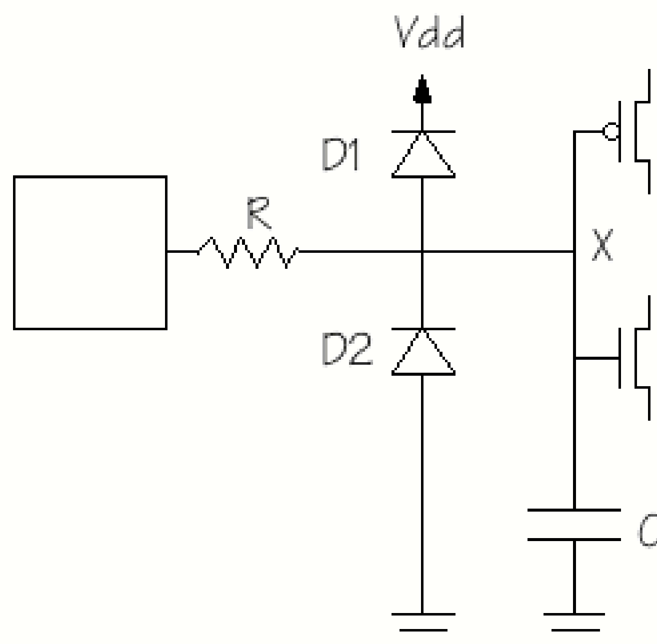
External outputs

- Need sufficient drive for required rise and fall time
- Add intermediate buffer stage to lower internal load
- Ratio of 2.7 is optimal for speed
- Two inverting stages yield a non-inverting output
- Susceptibility to latch-up is high
    - I/O currents are high
    - Excessive transients
    - Use guard rings tied to appropriate supply rails
- Driving TTL
    - Logic thresholds match OK
    - TTL low is 0.4V max, TTL high is 2.4V min
    - CMOS low is 0V, CMOS high is 5V
    - CMOS buffer must sink 1.6mA at < 0.4V



External inputs

- Static input protection
  - MOS transistor gate has high input resistance
  - Oxide breaks down at 40 to 100 volts
  - Diodes turn on when X rises about Vdd, below Vss
  - Resistor R limits peak current through diodes
  - Values of R range from 200 to 3000 ohms
  - R is a long diffusion or poly wire
  - Note RC time constant on input (delay!)
- TTL input
  - Set inverter switch point near 1.4 volts
  - Choice transistor size ratio to set switch point
  - Add pull-up resistor to pad to improve TTL high

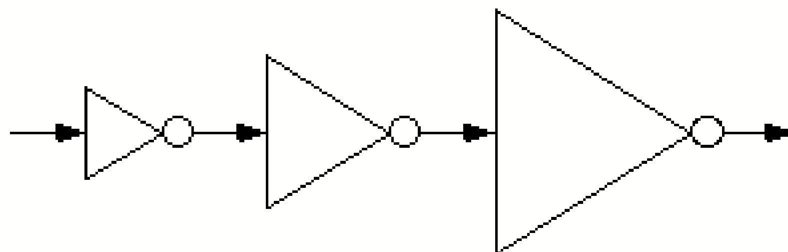


Clock distribution

- Chief problem: clock skew
  - Positive skew: clock arrives too late
  - Negative skew: clock arrives too late
  - Phase skew: clock phases early or late
- Why is skew a problem?
  - Clock pervades every part of chip design
  - Clock wires are the longest
  - RC constant can exceed delay of local logic
  - Clock must be driven into many cells (high fan-out)
- Techniques
  - Route on metal (RC constant small as possible)
  - Central clock driver
  - Distributed clock drivers
  - Cross-coupled driver design

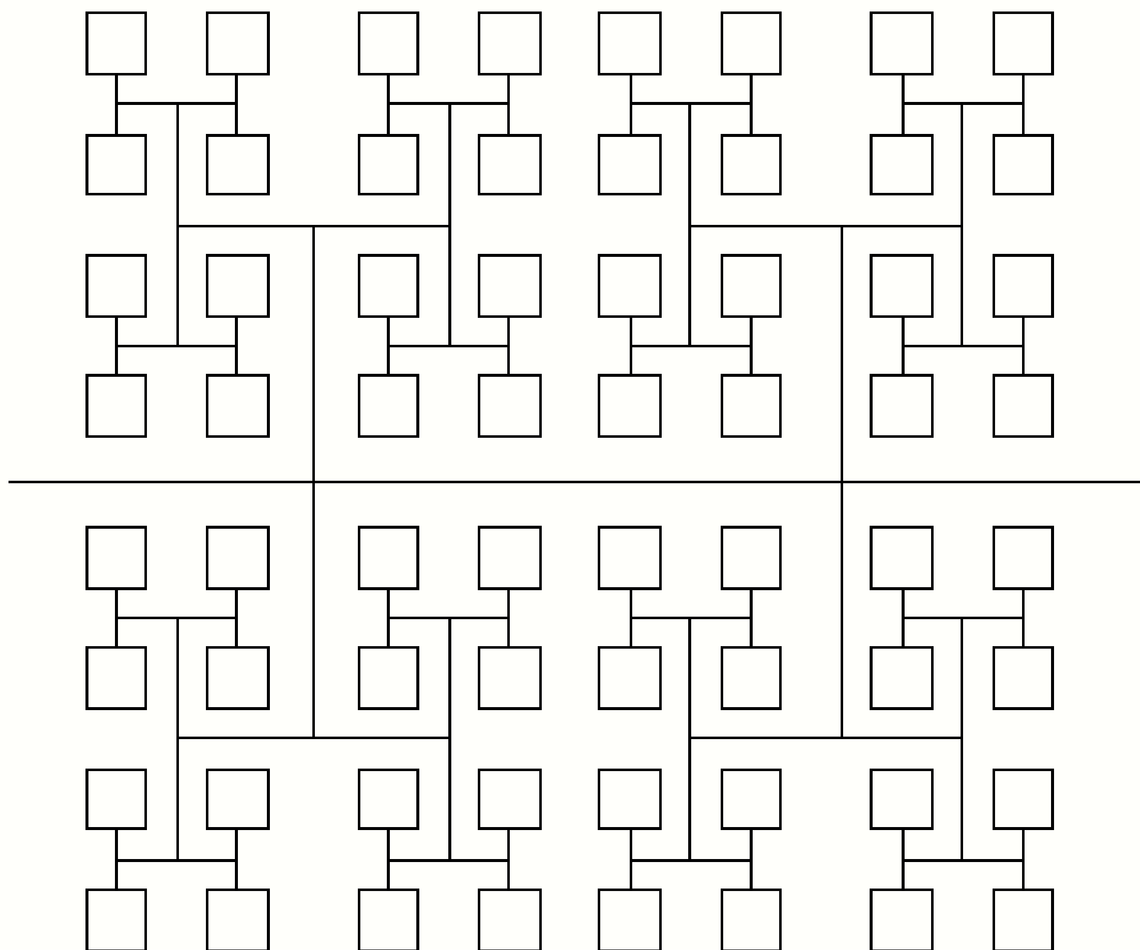
Central clock driver

- Drive all clock inputs from single point
- Use progressively larger buffer circuits
- Increase transistor width by four per stage



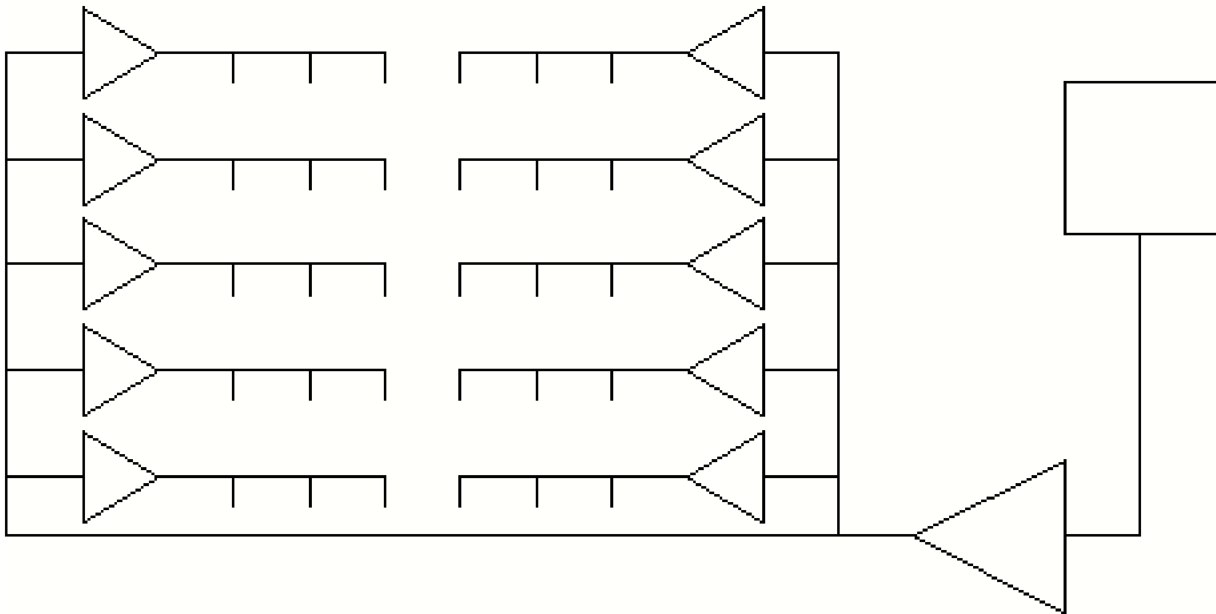
H-tree clock distribution network

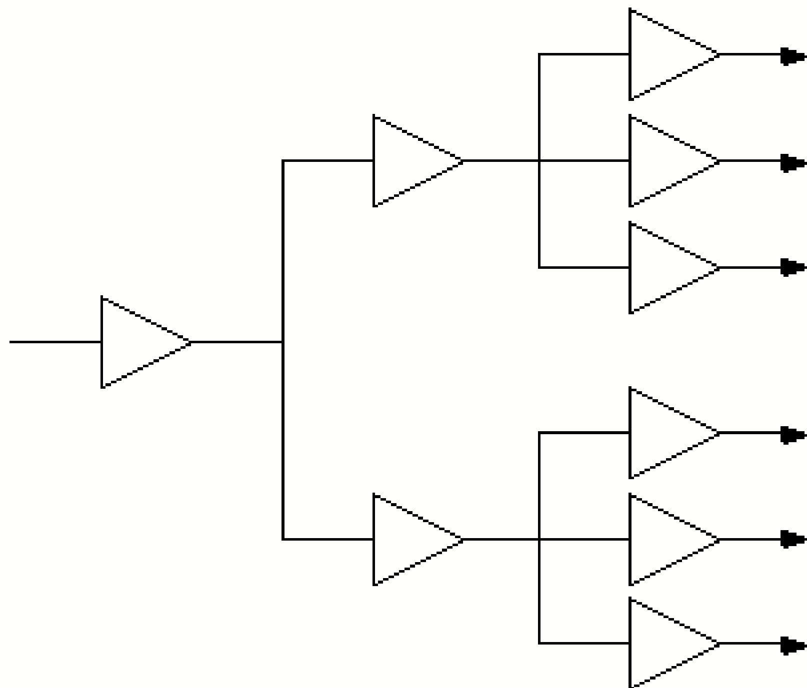
- Recursively layout tree in H's
- Load on the clock becomes quite large
- Steiner tree
  - Tree of minimum length to interconnect nodes
  - Minimize interconnect length
  - Minimize load
  - Need to add nodes to reduce length



Multi-level buffer network

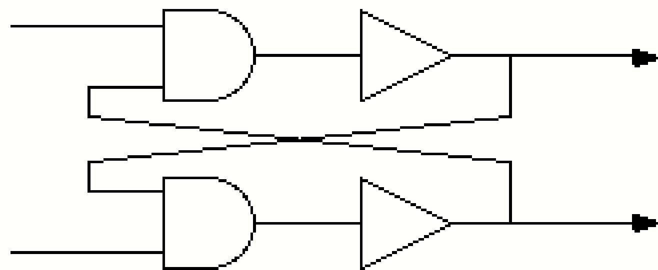
- Receive externally generated clock at input pad
- Drive and distribute clock to major subsystems
- Each subsystem buffers clock and distributes locally
- Relatively small drivers are distributed throughout





Phase skew

- Multi-phase clock (PHI1 and PHI2)
- Phases overlap (both PHI1 and PHI2 asserted)
- Causes
  - Different loading on PHI1 than PHI2
  - Assymetric circuit technology (longer rise than fall)
- Cross-couple drivers (add feedback)

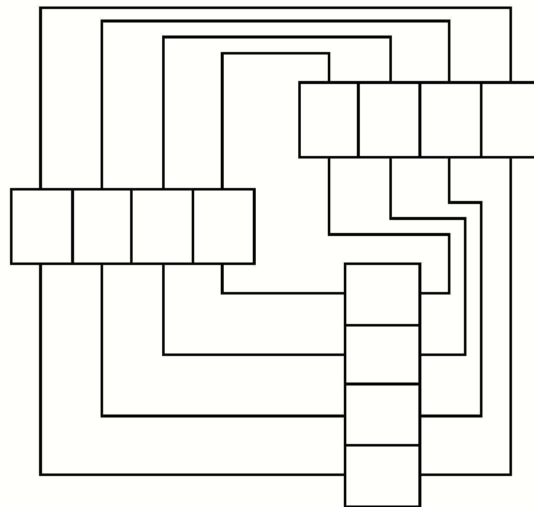


High speed clocks

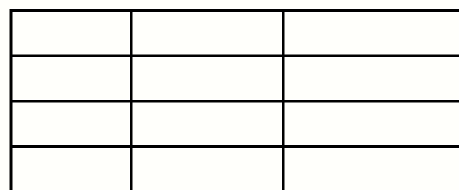
- A challenging design problem!
- 100 MHz clock, total clock period is 10 nsec
- Very short rise and fall times (about 0.5 nsec)
- Pulse duration in the 2 nsec range
- Driver
  - Very short rise and fall times
  - Must drive large capacitive load
- Analyze transmission line effects of clock network
- Technique
  - Equal series resistance on all clocks paths
  - Tune drivers to capacitive load on each line

Blocks versus bit-slices

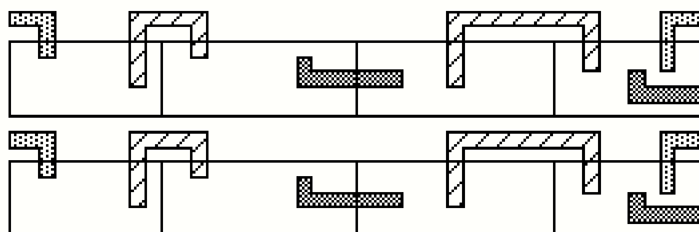
- Logically, we model systems using block diagrams
- Sees like a natural way to lay out a system
- Disadvantages
  - Too much random wiring and long busses
  - Space inefficient



- Pitch matching, abutment win for bit-sliced designs
- Even if cells are a little bigger, total area is less



- Leave wiring channels if necessary
- Useful for standard cell designs (e.g., ITD cells)



Feedback to the plan

- Layout is complete -- time to analyze
- Demonstrate that engineering constraints are met
- Speed
  - Recompute speed of critical path
  - Use real load capacitances
  - Simulate entire critical path or cells on path
  - $\text{Delay} = \sqrt{t_1^2 + t_2^2 + \ldots t_n^2}$
  - Evaluate drive and delay times on clock wires
- Space
  - Be sure design fits in payload area (frame)
  - Look for new opportunities for space reduction
  - Reconsider signal to pad assignment
- Current
  - Recompute subsystem current draw
  - Check wire sizing at each level of power grid
  - Is current density less than $1\ \text{mA} / \mu\text{m}$?
- Power dissipation
  - Compute static power dissipation
  - $P_{static} = \Sigma$ leakage current $\times$ supply voltage
  - Leakage current is $0.1\ \text{nA}$ to $0.5\ \text{nA}$ per gate
  - Find lumped capacitive load of chip (sum of loads)
  - Estimate dynamic dissipation using total load
  - $P_{dynamic} = C_L V_{dd}^2 f$
  - $P_{total} = P_{static} + P_{dynamic}$
  - Is total power dissipation within max for package?

Material

- Capacitance ($\lambda = 1.5\ \mu m$)
  - Unit is capacitance per square micron of area
  - Compute area and multiply by material constant

Gate	4.5×10^{-4}	
Polysilicon over field	0.5×10^{-4}	
n-diffusion (active)	0.9×10^{-4}	$pF / \mu m^2$
p-diffusion (active)	0.9×10^{-4}	
Metal 1 over field	0.2×10^{-4}	
Metal 2 over field	0.1×10^{-4}	

- Resistance
  - Sheet resistance
  - Count number of squares of material
  - Multiply by material constant
  - Resistance $R = \frac{\rho}{t} \frac{L}{W}$

ρ = Resistivity
t = Thickness
L = Conductor length
W = Conductor width

Metal	0.5 Ohms / square
Silicides	3 Ohms / square
Diffusion	25 Ohms / square
Polysilicon	50 Ohms / square

Capacitance

- Charge storage
- Wires and gates are all capacitors
- Charge on a capacitor

$$Q = C \times V$$

Q is charge, C is capacitance, V is voltage

- Capacitance

$$C = \frac{KA}{d}$$

K is dielectric constant, A is plate area, d is distance

- d and K are given for a particular process
- Capacitance is stated as farads per square-micron
- Precise estimation should account for
  - Area exposed to bulk
  - Side wall exposed to field
  - Side wall exposed to gate region
- Compute area and multiply by material constant
- Typical capacitances by material

Material	Capacitance	Unit
Gate	4.5×10^{-4}	$pF / \mu m^2$
Polysilicon over field	0.5×10^{-4}	
n-diffusion (active)	0.9×10^{-4}	
p-diffusion (active)	0.9×10^{-4}	
Metal 1 over field	0.2×10^{-4}	
Metal 2 over field	0.1×10^{-4}	

- Keep wires short and route on metal if possible
- Use diffusion and poly for short local wires only
- Propagation depends upon the RC time constant

Resistance

- Resist the flow of current
- Used as a current limiting device
- Ohm's law

$$V = I \times R$$

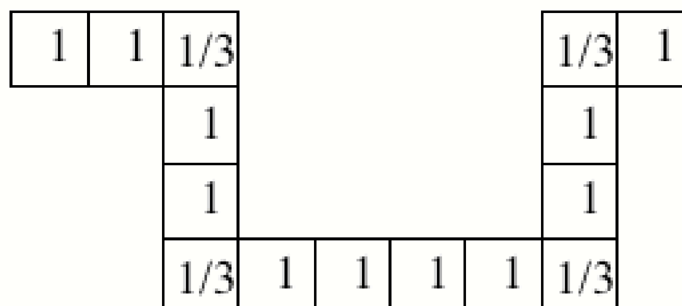
V is voltage, I is current, R is resistance

- Resistance

$$R = \frac{\rho}{t}\frac{L}{W}$$

ρ is resistivity, t is thickness, L is length, W is width

- Sheet resistance
  - t is given for a particular fabrication process
  - Resistance is stated as ohms per square
  - Estimation
    - Break shape into squares
    - Count number of squares of material
    - Multiply by material constant
    - Corners count as 1/3 of a square



Resistance (2)

- Typical sheet resistances

Metal	0.5 Ohms / square
Silicides	3 Ohms / square
Diffusion	25 Ohms / square
Polysilicon	50 Ohms / square

- Routing priority
  - Ground rail
  - Clock lines and time critical signals
  - Positive power rail (Vdd)
- Rules of thumb
  - Route power and ground on metal
  - Route signals on metal if possible
  - Polysilicon and diffusion over short distances
  - Use metal for long distances
  - Keep wires short
  - Plan ahead for shortest wire routing
- Crossovers
  - May be necessary to route under metal
  - Use contacts to change layers
  - Metal may cross poly, n-diffusion or p-diffusion
  - Favor polysilicon for signal lines
  - Use diffusion for power connections