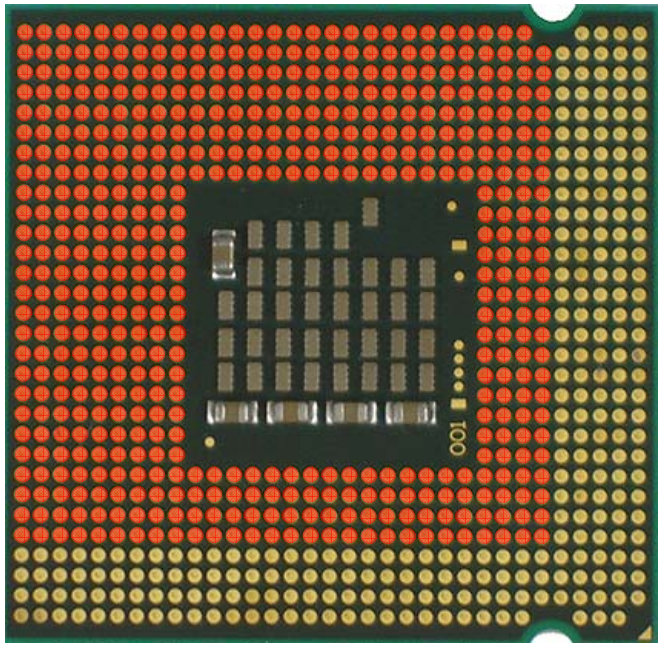
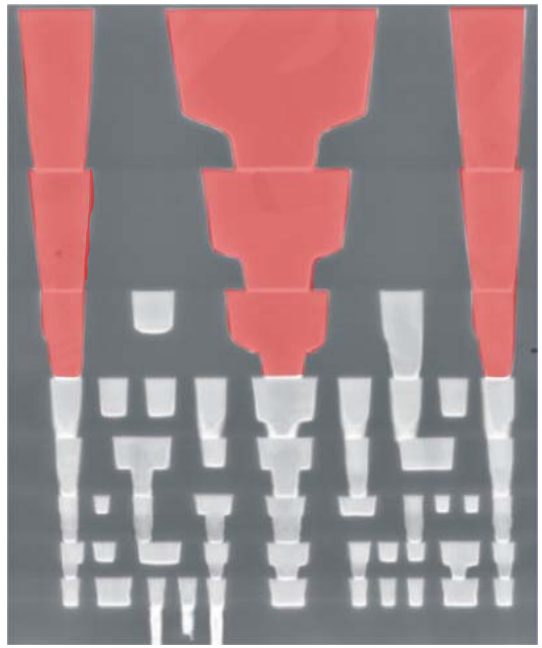
Lecture 17

Low Power Circuits and Power Delivery

Computer Systems Laboratory
Stanford University
horowitz@stanford.edu

Power Delivery Is Resource Intensive

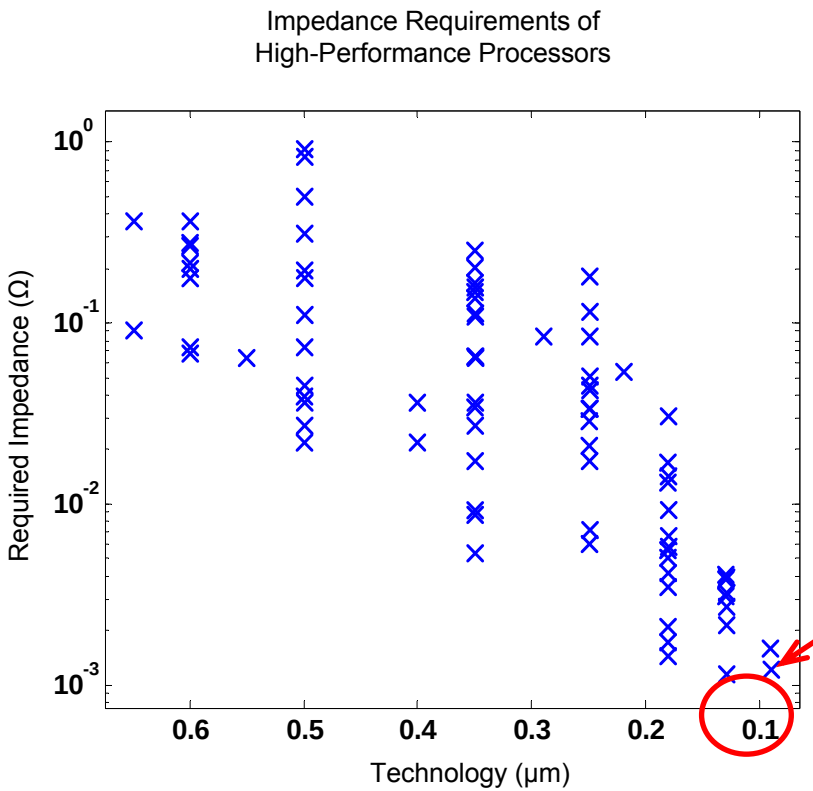




- Significant time and resources spent on power distribution network:
  - ~70% of package pins just for power
  - Top 2-3 (thick) metal layers
- Why has power delivery become this critical?

Scaling and Supply Impedance

- CMOS scaling has led to lower supply voltages
  - With constant (or increasing) power consumption



- This forces drastic drop in supply impedance
  - Even at constant power:
  - $V_{dd} \downarrow$, $I_{dd} \uparrow \rightarrow |Z_{required}| \downarrow\downarrow$
- Today's chips:
  - $|Z_{required}| \approx 1\ m\Omega$!
- Hard to achieve across entire frequency spectrum
  - Supply voltage will be noisy

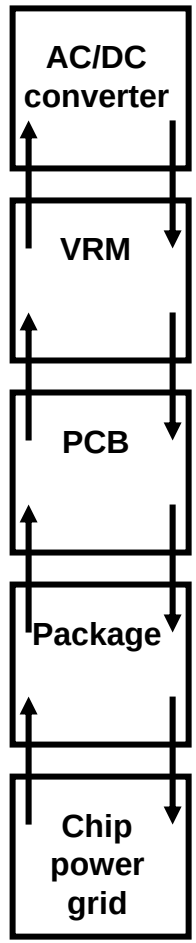
Power To The Chips

- Today's microprocessors pushing 100s of amps
  - Itanium: 1.2V, 130W
  - Opteron: 1.2V, 95W
  - Not all: PentiumM uses 20A, ULV 1GHz Celeron uses 5A
- Tomorrow's supercomputers have a 10MW limit
  - At a power supply of, say, 1V, that's a lot of juice
- Okay, you might not be building supercomputers
  - But you will still need to push in lots of amps into your chips
- What are the designs and tradeoffs involved in power networks?

Power Distribution Network

faster response

- AC/DC converter
  - Usually 110VAC to 12 or 5VDC in desktop PCs
- Voltage Regulator Module
  - Converts one DC level to another (5V to 1.2V)
- Printed circuit board
  - Planes send current from VRM to the package
  - Planes have capacitance for bypass; use discretes too
- Package
  - Deliver current to the chip itself using balls or bonds
  - Can use bypass caps on the package as well
- Chip power grid
  - Use device bypass capacitors



Power Supply Goals

- All levels: Provide power to the chip transistors
  - Maintain the voltage during chip operation (i*R noise)
    - Wide traces on-chip; thick copper in PCB (1 “ounce” Cu = 35μm thick)
  - Maintain the voltage during switching transients (Ldi/dt noise)
    - Sufficient bypass capacitance throughout the path

oz/sq-ft

- On-chip: Shield and stabilize signal wires on the chip
  - Isolate sensitive signals, like clocks, to prevent coupling
  - Provide current return paths for signals that doesn’t impact R_{total}

- On-chip: Consume minimal area, design time, wire tracks
  - Avoid electromigration problems from too-narrow wires
  - Ex: Alpha 21064 used power planes (min design time, max area)

Chip DC current requirements

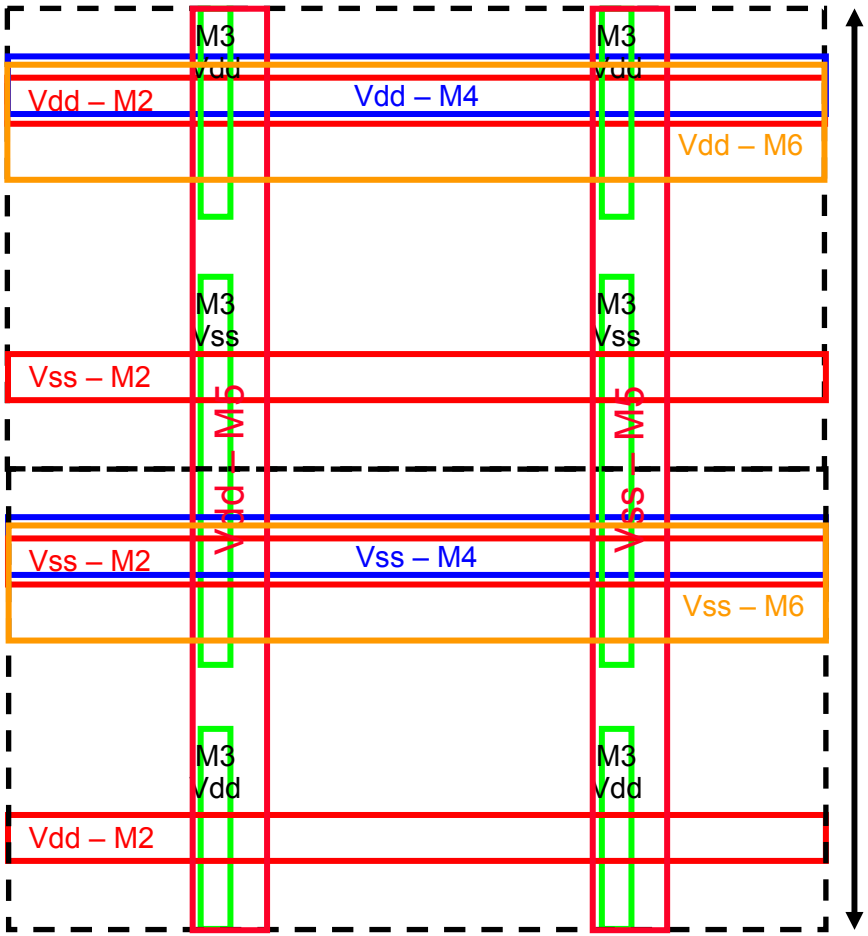
- Chip power supply designs exploit *regularity*
  - Top layers of metal use a strictly defined template, for example

  - Guarantees a minimal metal coverage for Gnd, Vdd, Vdd2, Vdd3…

- Vias require some extra care
  - Overlap metal power (M5 under M7, M4 under M6) to stack vias
    - Straight shot from M1 to Mtop is ideal, if you can line up the vias
    - Although newer technologies don't let you stack vias too high
  - Vias are generally Cu now – much better resistivity than W
    - Approximately 1Ω per via in Copper, 5Ω per via in W

- Appropriate templates can give low total "number of squares"
  - Good for DC voltage

6 Layer Power Grid Example – CBD



- Representative power grid design for 6 layer CBD shown
  - Custom layout may not be as regular at M2 & M3
- M2 is mirrored for well abutment
- M3 power shares tracks to limit metal usage and increase via counts
- Vias located at all next layer crossings
- Power metals are stacked as much as practical to simplify via stacks

Source: Ayers, Intel

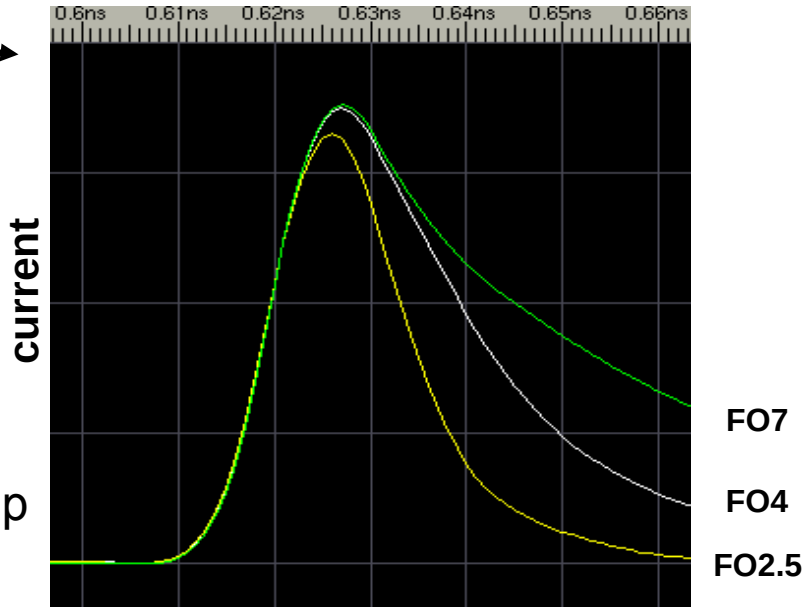
Chip AC Current Requirements

- A quick note on terminology
  - "AC" = switching events that generate high-freq noise
  - "DC" = constant current that causes i*R drop

- AC frequencies related to, but not equal to, clock frequency
  - They arise from the edge rate of signals on the chip
    - Knee of the frequency components curve: $(2\pi * T_{rise})^{-1}$
  - Fast edge rates generate high-frequency events and noise
    - Regardless of the clock frequency
    - Slowing the chip down doesn't reduce noise (maybe the sensitivity)
  - Slew-rate control common on off-chip I/O
    - Slow down the edges
    - Reduces the injected noise without much increase in latency

Transistor Switching Noise

- How much current does a switching inverter require?
  - 90nm simulations
  - Charge = area under the curve
  - Q = C*ΔV sets the required cap
    - ΔV is the maximum droop



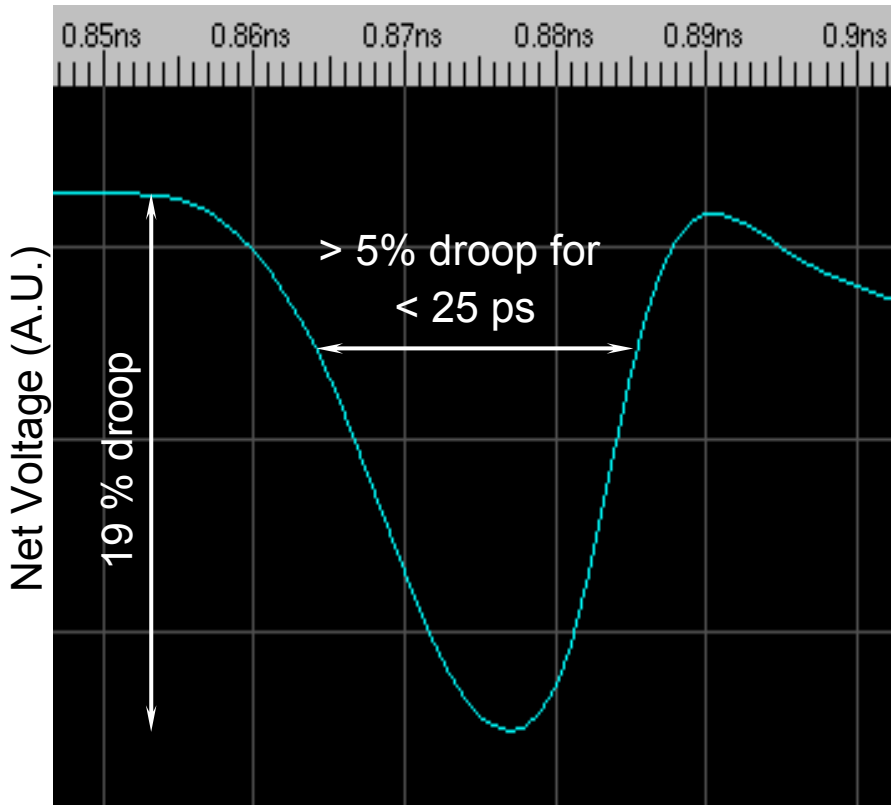
Source: Ayers, Intel

- Worst points for DC and AC?
  - DC: Worst i*R drop at peak
  - AC: Worst Ldi/dt midway up ramp
    - Fast: It's all over in 20pS

- Load doesn't matter (trailing edge is slow)
  - Driver size is important

Impact on Nearby Logic

- Very fast transistor switching means very fast noise spikes
- Random block of logic is usually not a big noise concern
  - Thousands of scattered small transistors fire at various times in a clock cycle
    - Not enough microns of transistor firing at once to cause a serious disturbance
  - Bad case will usually be a bank of synchronous drivers (like repeaters)
    - 64-256 large drivers firing synchronously
- Wave shown is from a power model repeater bank simulation with 90 nm technology
  - Spike droops up to 19% of Vdd
  - But droop only exceeds 5% of Vdd for < 25 ps

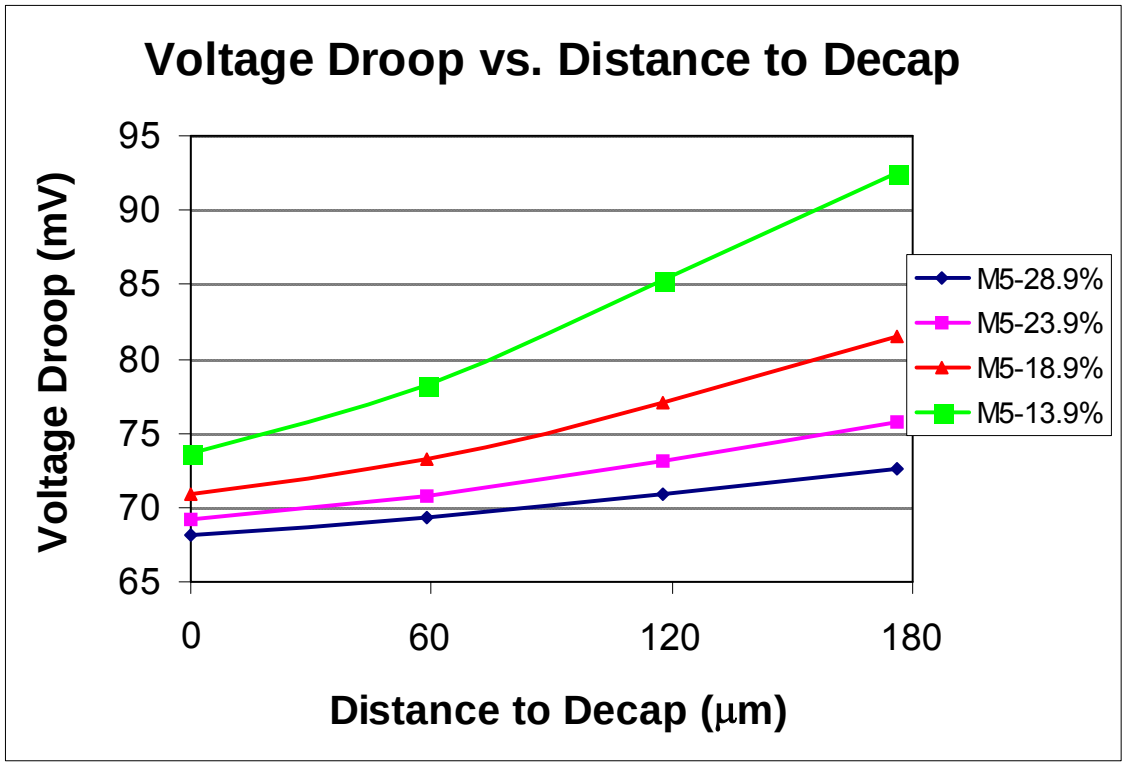


- With a clock cycle > 200 ps, there is minimal delay impact to nearby logic from one spike
  - Is extra decoupling really needed?
  - Noise spikes have the greatest speed impact on the repeated signal itself

Source: Ayers, Intel

Droop vs. Decap Distance and Die Metal

- Simulations from 180 nm technology node
  - Capacitors placed at various distances from noise source
- Note noise increase as capacitors are placed further away
- Substantial improvement with increasing power metal use



Source: Ayers, Intel

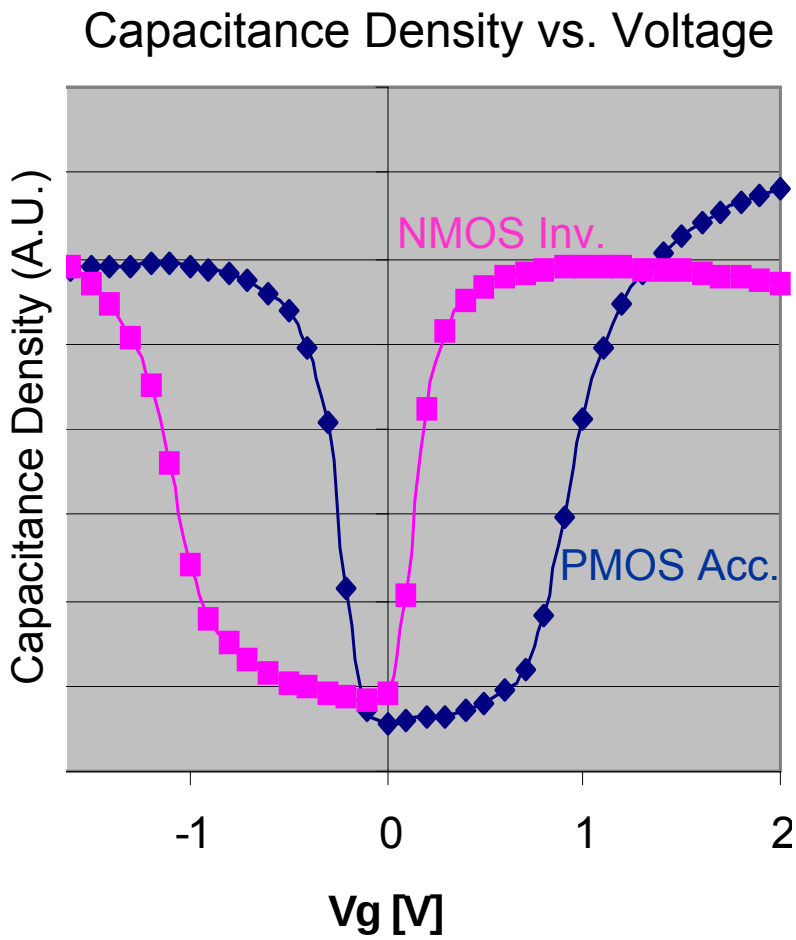
On-Chip Bypass Capacitance

- There is lots of vdd-gnd capacitance on a chip
  - Wire bypass cap: Vdd and Gnd wires can be near each other
  - "Natural" bypass cap
    - At any given moment, most gates are not switching (esp. memories)
  - Intentional bypass cap: inserted by the designers
    - This cap dominates; > 80% of total bypass from bypass cells
    - Terminology: "Decaps" = **de**coupling (bypass) **cap**acitor**s**

- Make bypass capacitors out of gates
  - For large capacitance (good), make W and L both very large
  - But for low resistance (good), make L relatively small, around 10λ
  - Gate oxide is thin, so a gate has a high capacitance density (good)
  - Gate oxide is thin, so the gate leaks current (bad)

Decoupling Capacitor Design (Cont'd)

- Cell type can be important
  - NMOS faster than PMOS inversion cells
  - PMOS accumulation cells can be faster than inversion but require wells which eat up space
  - Gate oxide leakage concerns may force accumulation cells
    - Work function shift reduces leakage
    - But capacitance rolls off at lower voltages (see graph)
    - Not well suited for analog circuit applications



Source: Ayers, Intel

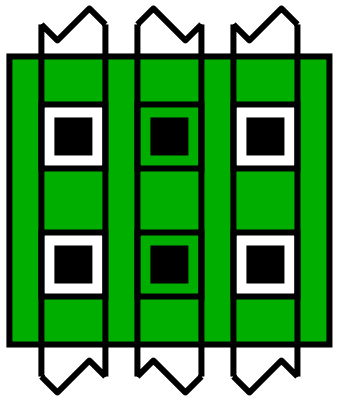
Fill Cells

- Typical method is to use decaps as fill blocks
  - Chips are never *completely* full of transistors
  - We often open up wiring channels between blocks
  - Must fill these wiring channels
    - Need to route the required wires
    - Need to fill metal on the other layers to hit minimum density rules (30%)
    - Can opportunistically fill these channels with bypass decaps
    - Also helps with required poly density across the die (15%)

- Fill cell decaps should be big and widely spaced for yield
  - Tie them into the power grid directly as a repeatable layout cell
  - Remember to go back and modify your schematic
    - They are devices, after all
    - They will affect your LVS (layout vs. schematic) checks

What Decap Cells Are Useful?

- Draw a “waffle”-style decap
  - Here, inversion decap shown; accumulation decap analogous



A sheet of poly (green) that rests over inversion charge

Four holes cut out for Gnd connected diffusion
Diffusion mostly there to provide the inversion layer charge

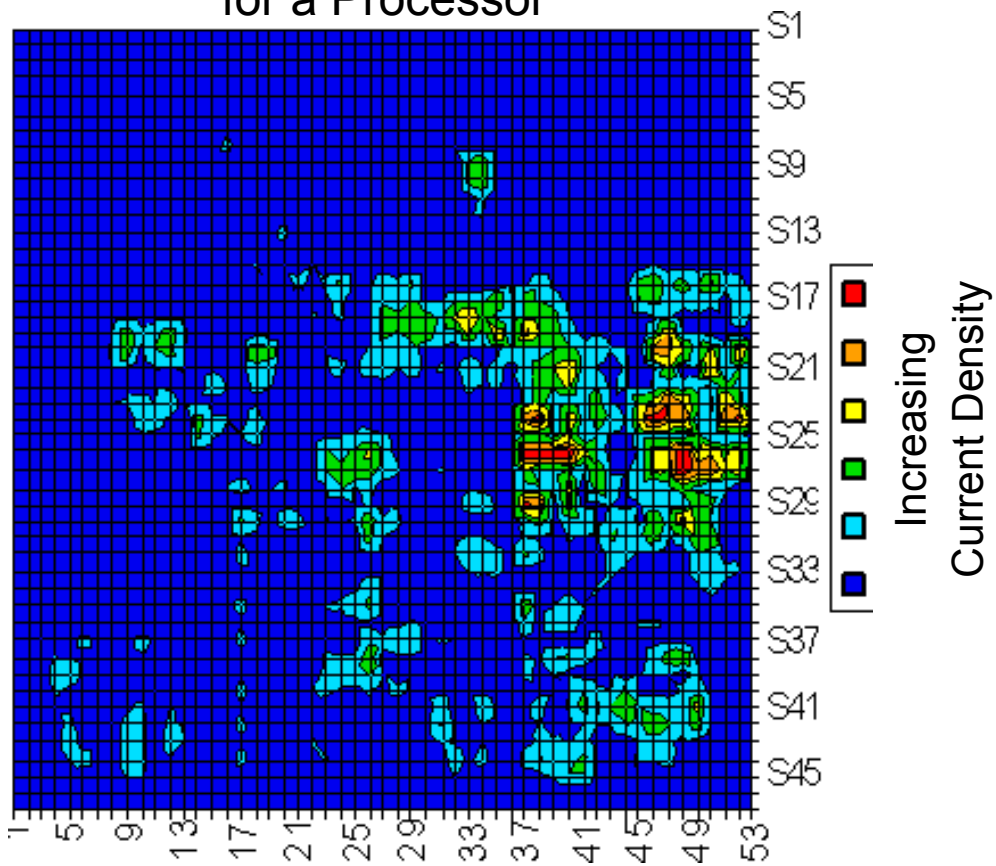
Poly connection to M1 happens in the middle stripe

- Don’t place decaps too far from areas of high current change
  - Current must travel from decap to areas of use
  - Only decaps within 100-200μm of circuits are useful
- Problem is you need to worry about fill rules …

Moving Up from Chip: Package Connection

- C4 bump pitch has not been scaling as fast as transistor technology while current density is scaling
  - Result is increasing current per bump which will stretch reliability limits
- Note that only a few small areas have the highest current
  - Technology and uarch solutions are likely to be needed
- Increased top and second layer metal resources will also be needed

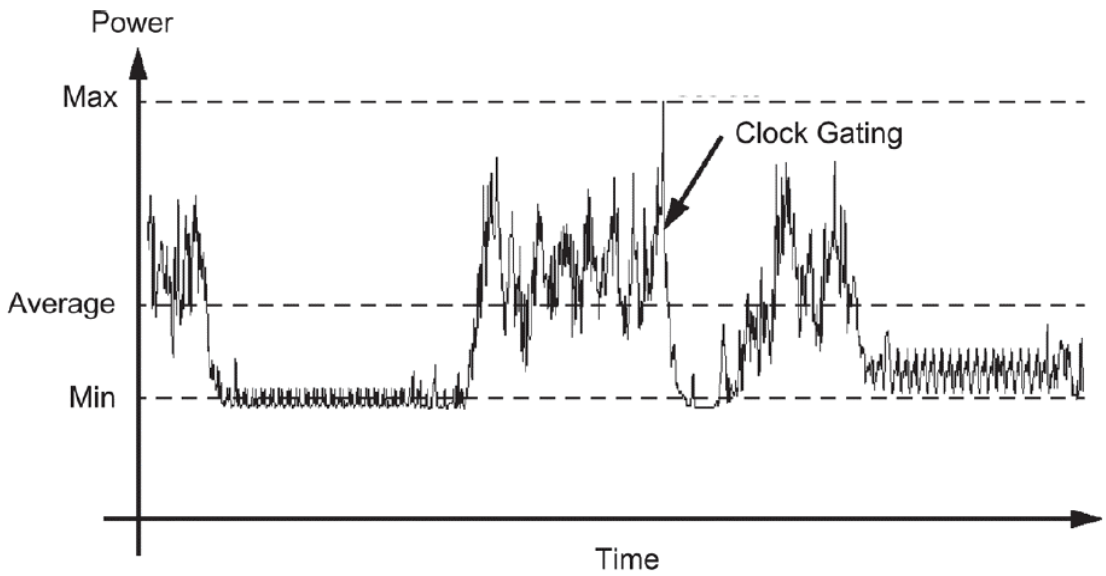
C4 Bump Current Density for a Processor



Source: Ayers, Intel

di/dt: Current vs. Time

- Example profile of current during chip operation
  - Full-chip circuit switching events summed together
  - Sun Microsystems CPU simulation



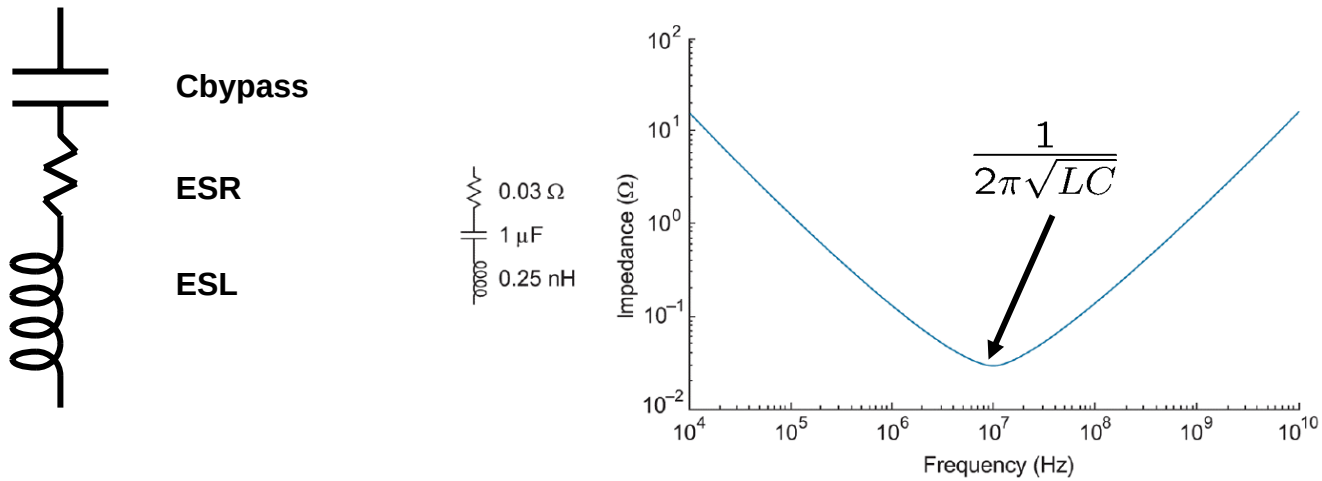
Source: Harris, Addison-Wesley '05

  - Many low power techniques make di/dt worse

Bypass Cap Frequency Response

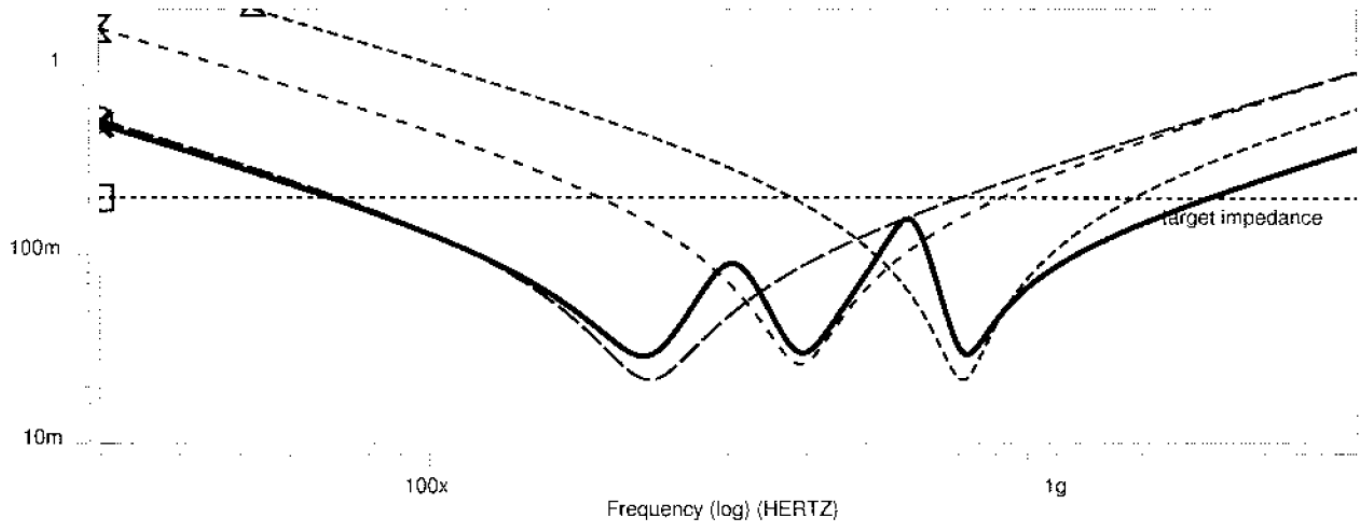
- Every bypass capacitor has some parasitics
  - Equivalent series resistance (ESR) and inductance (ESL)

- Frequency response
  - At low frequencies, we get a high impedance (Cbypass)
  - At high frequencies, we get a high impedance (ESL)
  - Somewhere in the middle we get a pure resistance (resonance)



Meeting Target Impedance

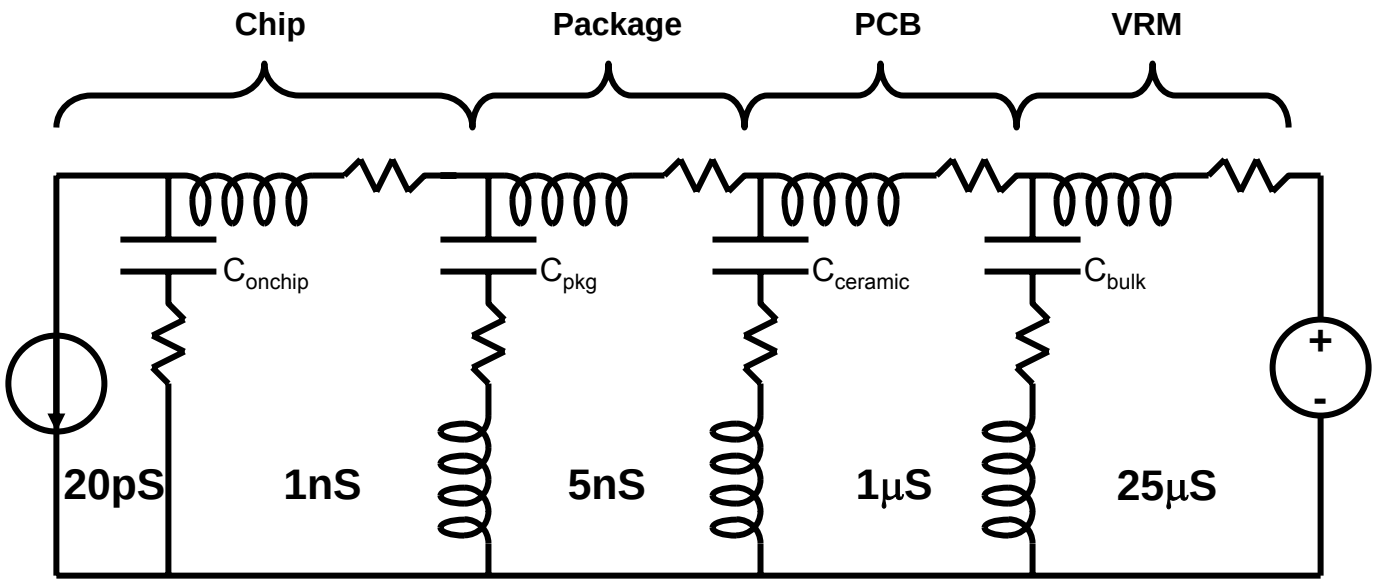
- So we just have to add capacitors until we've hit our target
  - Bulk capacitors good to keep down impedance at low frequencies
  - Ceramic capacitors near the package good at mid frequencies
  - PCB/package capacitors next to die extend to higher frequencies



Source: Smith, TransAdvPack, '99

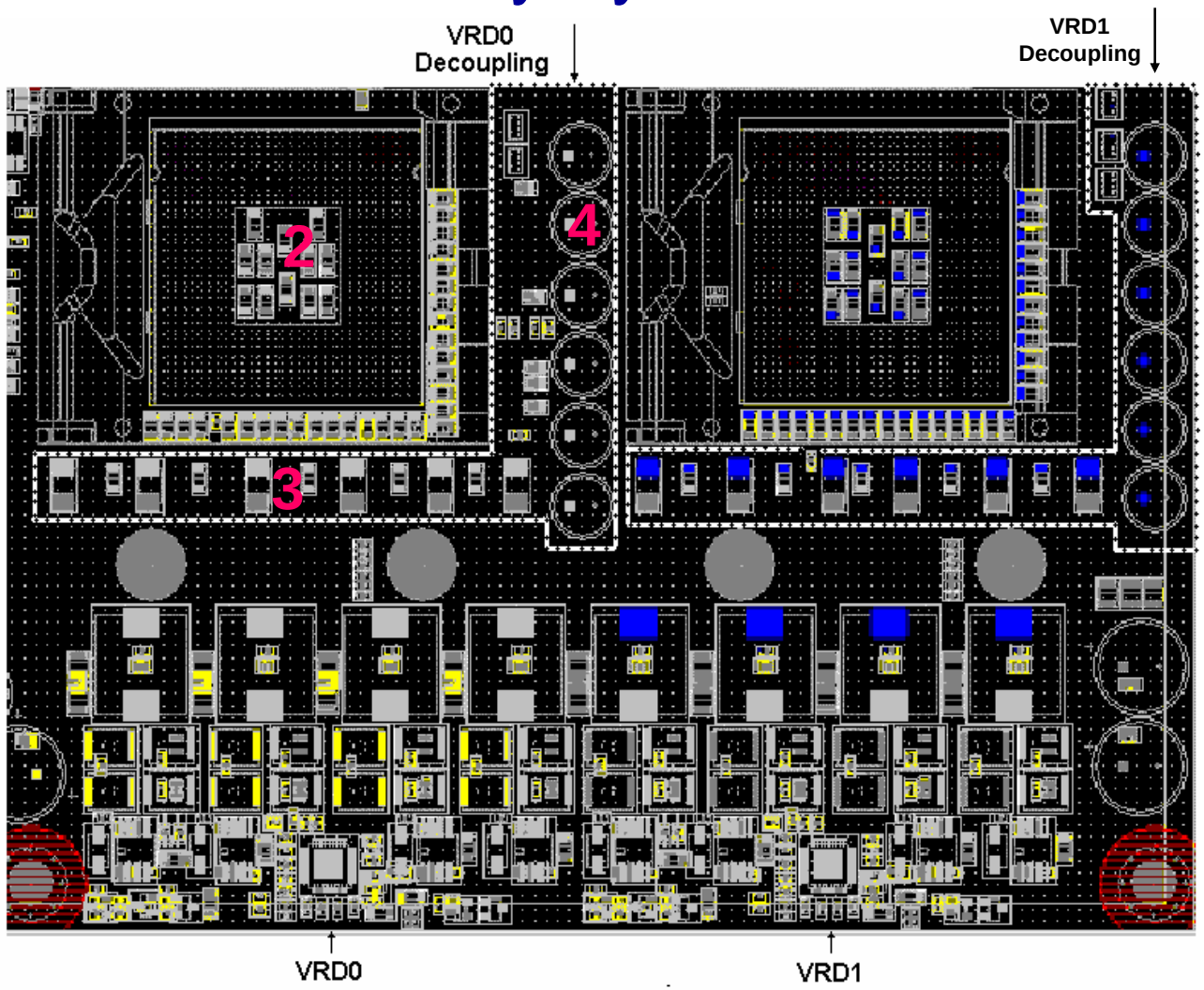
Bypass Capacitance

- Switching events are far too fast to pull current from far away
  - VRM can respond only in ~25μS; 3 orders of magnitude too slow
- Feed current from more local sources using bypass capacitors
  - Capacitors act like (imperfect) batteries



Typical Power Delivery System

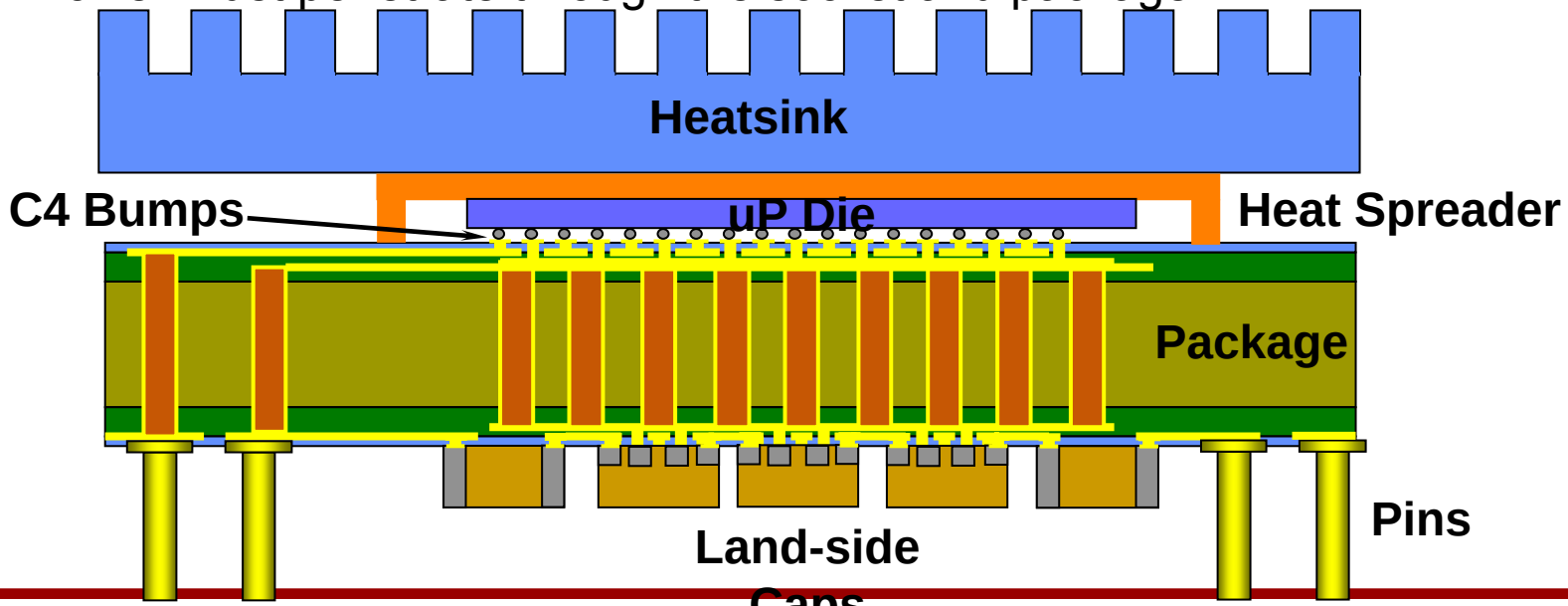
- 2 processor MB design shown
- Voltage Regulators are located close to processors
- VR current brought in to processors on ~2 sides to reduce impedance
- Note the levels of decoupling
  1. Die (MOS)
  2. Back of package
  3. High speed MB
  4. Low speed MB



Source: Ayers, Intel

Packaging Cross-Section

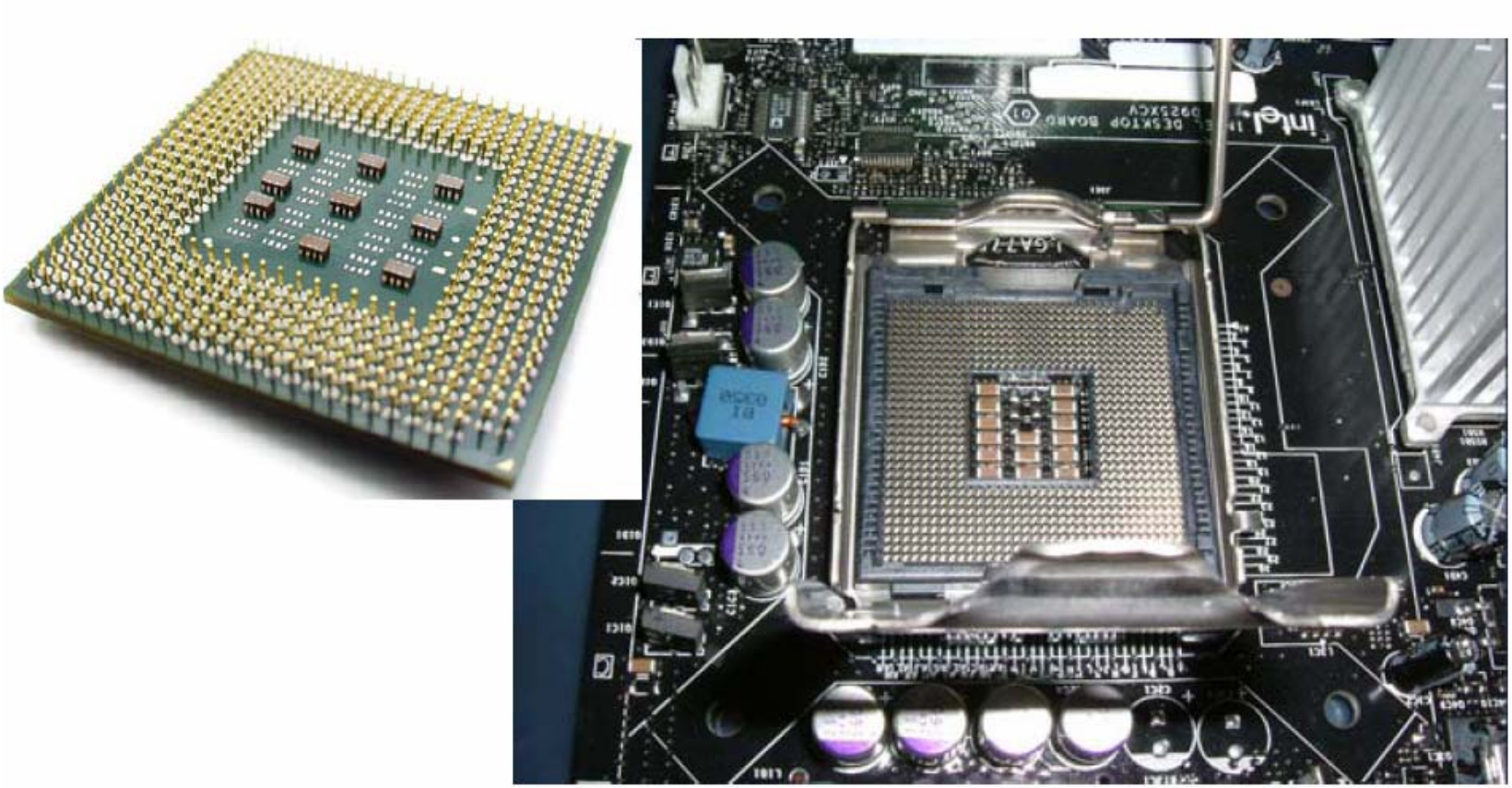
- A sample processor cross-section is shown below
  - May or may not have a heat spreader
  - May have die side capacitors as well as land side
  - Package may have 4-14 layers depending on number of signals and cost structure of market (low-end desktop to high-end server)
  - May have an additional layer of package (interposer) for space transformation and for housing additional components
- Power must penetrate through the socket and package



Source: Ayers, Intel

Bypass Capacitances in Real Life

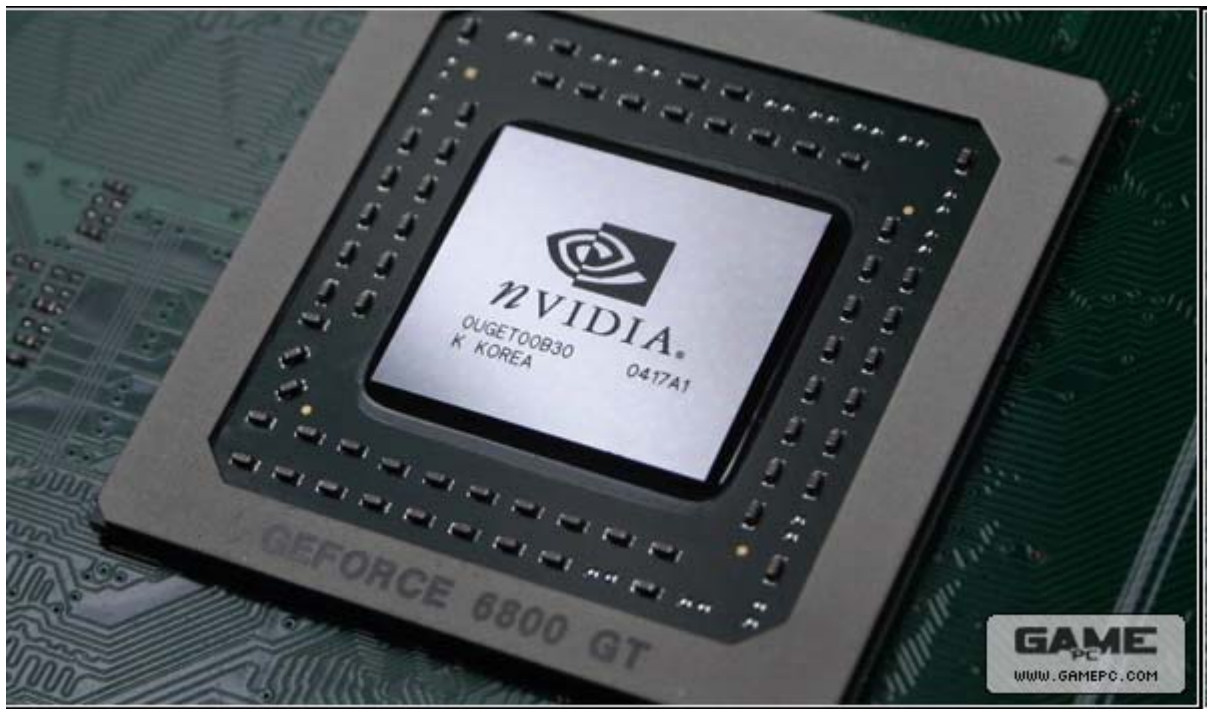
- Left: package bypass; Right: PCB bypass



Source: Mai, CMU

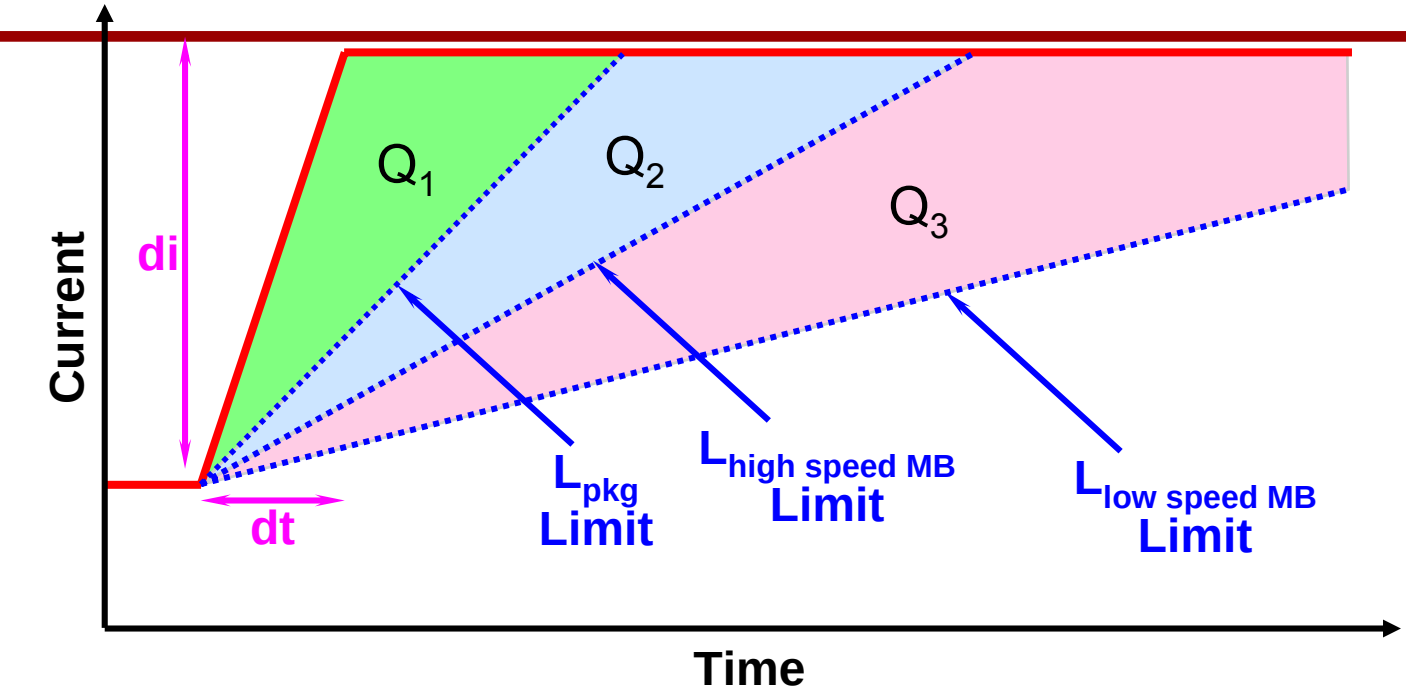
More Bypass Capacitors

- NV40 GPU



Source: gamepc.com

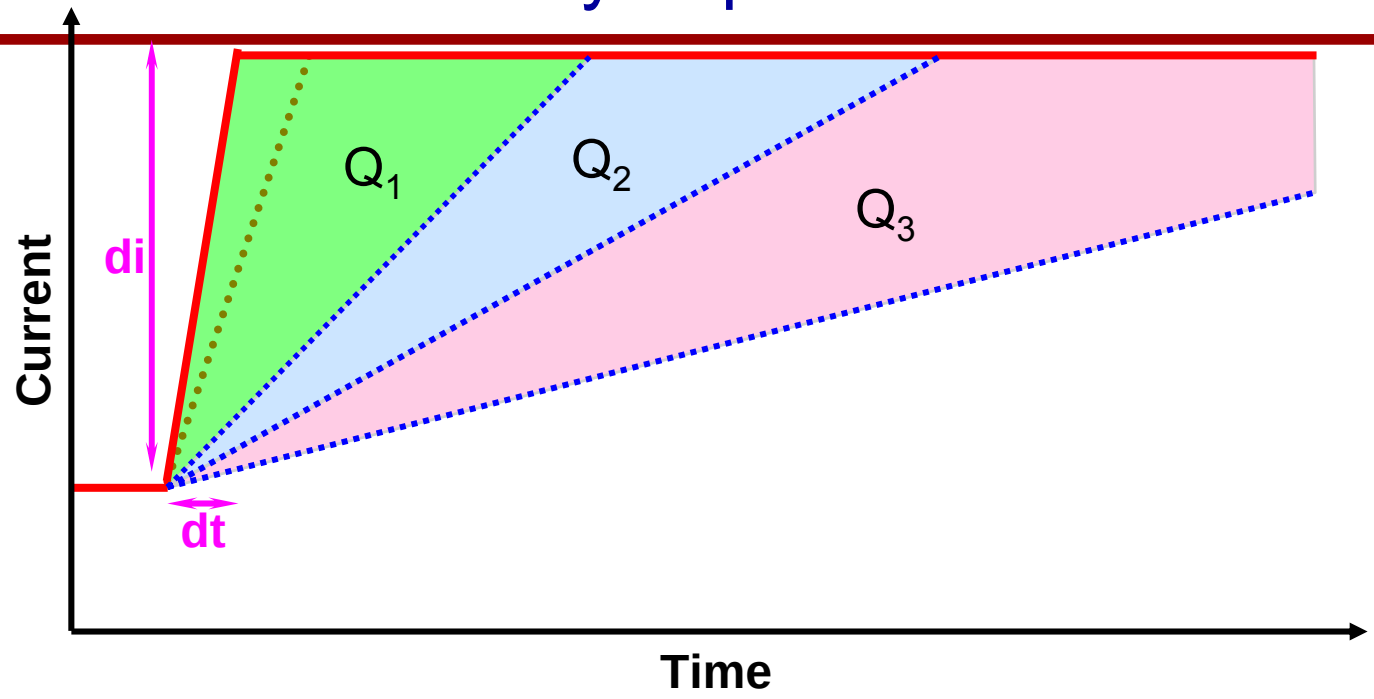
Factors in Determining Decoupling



- The area of triangle Q_1 determines the need for die capacitance
  - $C_{die} = Q_1 / \Delta V$; determined by di, dt, L_{pkg}, and the voltage drop target
- The area of triangle Q_2 determines the need for package capacitance
  - $C_{pkg} = Q_2 / \Delta V$; determined by di, L_{pkg}, L_{HSMB}, and the voltage drop target
- The area of triangle Q_3 determines the need for board capacitance
  - $C_{board} = Q_3 / \Delta V$; determined by di, L_{HSMB}, L_{LSMB}, and the voltage drop target

Source: Ayers, Intel

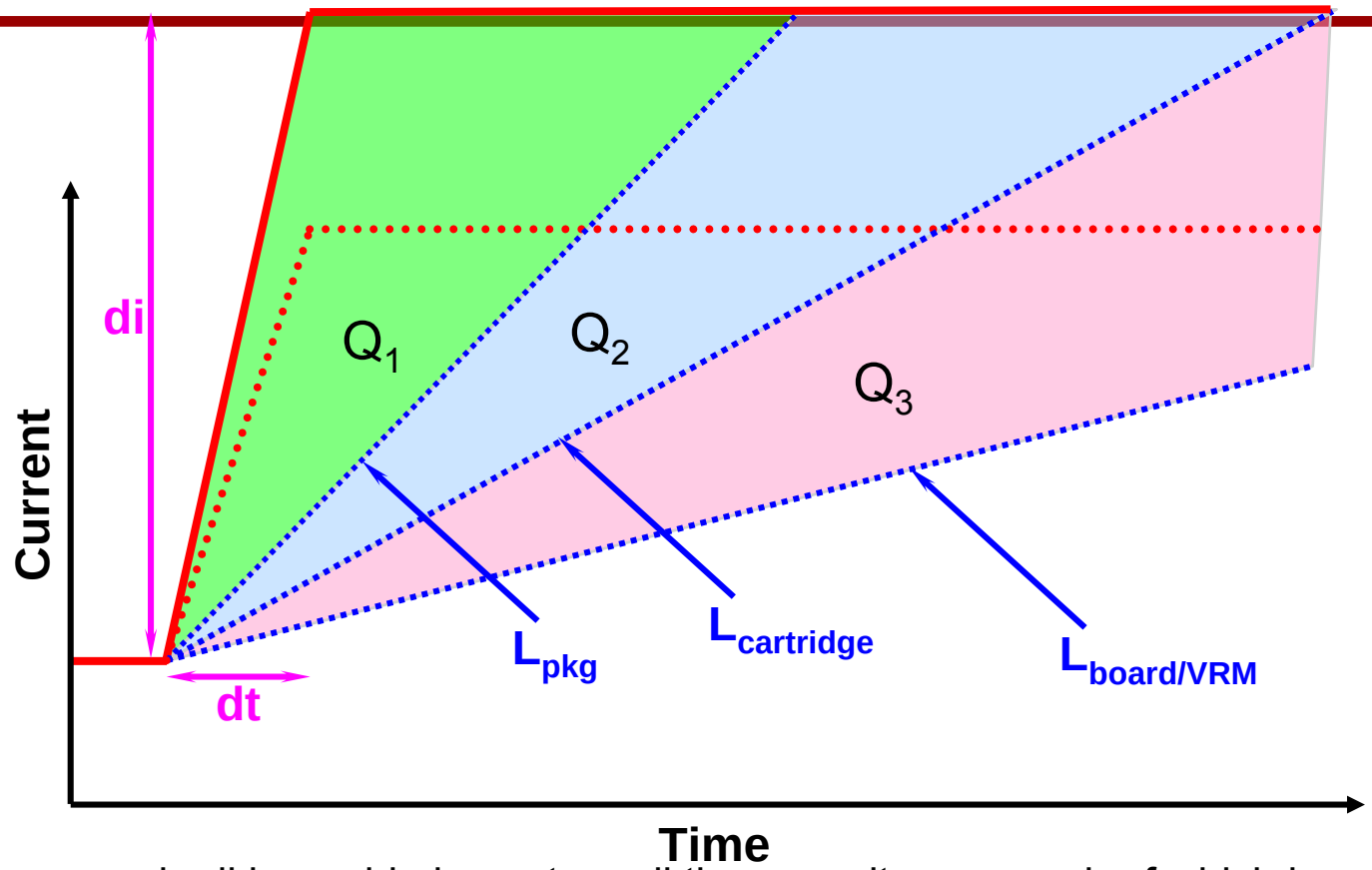
Power Delivery Implications – dt



- Picture shows dt decreased by 2x from previous page -- small impact
- Capacitances are proportional to triangle areas
  - Note that the area of the Q_1 triangle (die capacitance) increases by less than 2x
  - Area of the other triangles (other capacitors) are unaffected

Source: Ayers, Intel

Power Delivery Implications – Imax

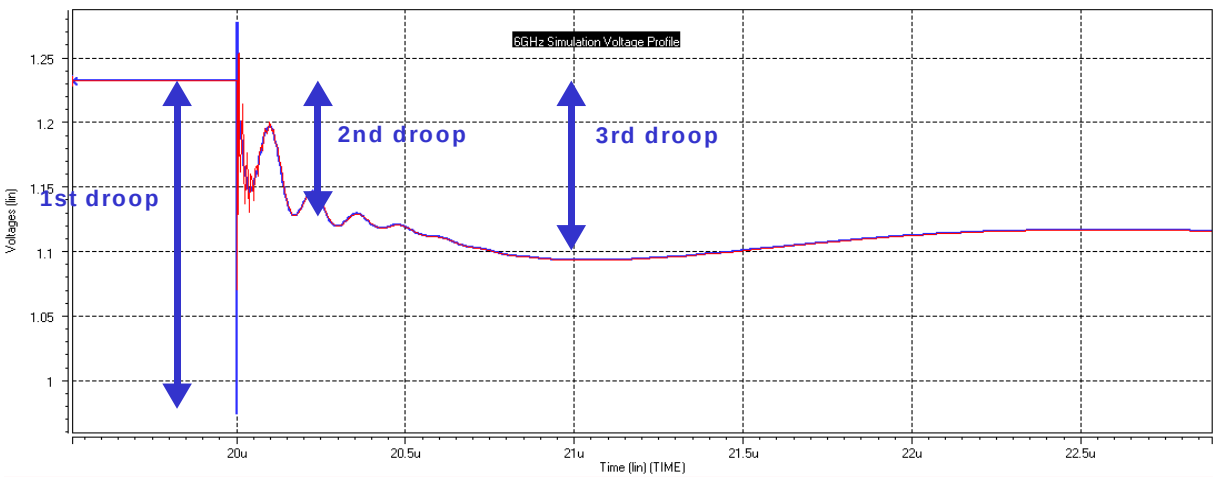


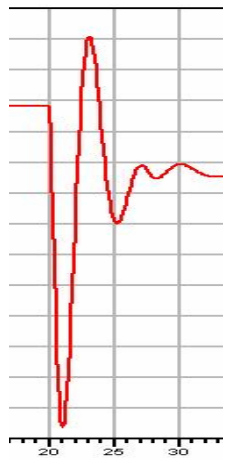
- An increase in di has a big impact on all the capacitances each of which is proportional to the triangle areas
  - Square relation for area: 2x increase in di increases the triangles by 4x!
  - Even greater increase for Q_1
- Reducing di is most effective for voltage control

Source: Ayers, Intel

Step Response

- Voltage response for a complete power delivery system
  - Simulated response
  - Each droop happens when a new bypass cap kicks in

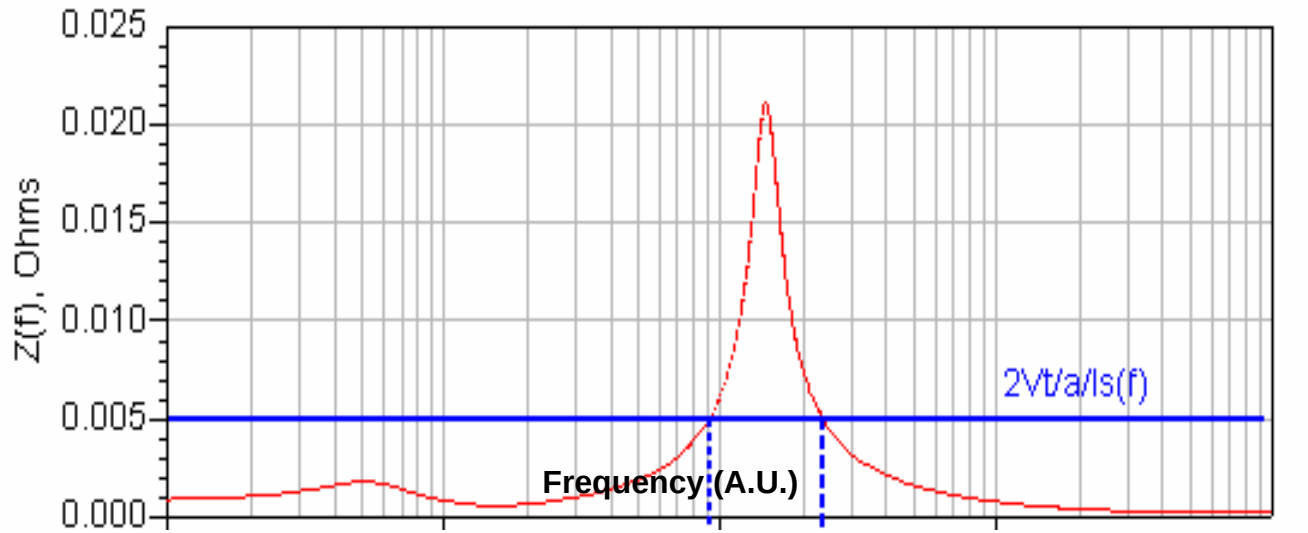




1st Droop Zoom In

Source: Ayers, Intel

Frequency Domain System Modeling

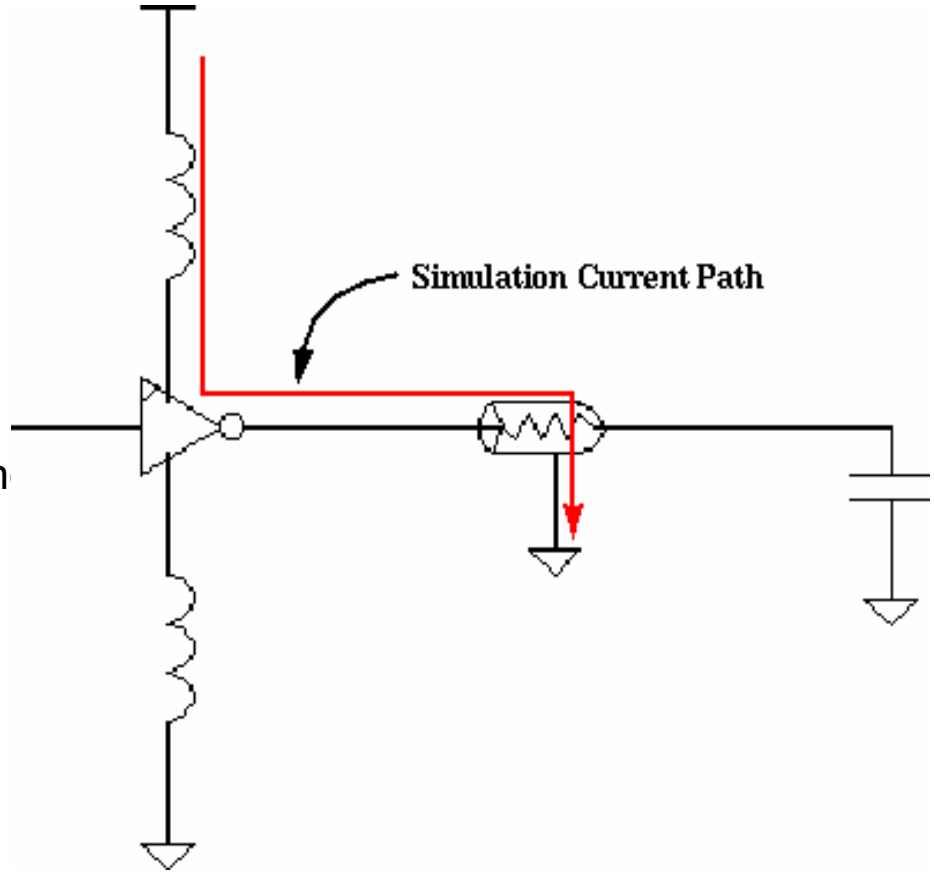


Source: Ayers, Intel

- Take transform of impulse response
  - Get the impedance vs. frequency
- Ideal this would be a flat line
  - Has peaks due to resonance
  - Worst peak is package inductance / chip capacitance

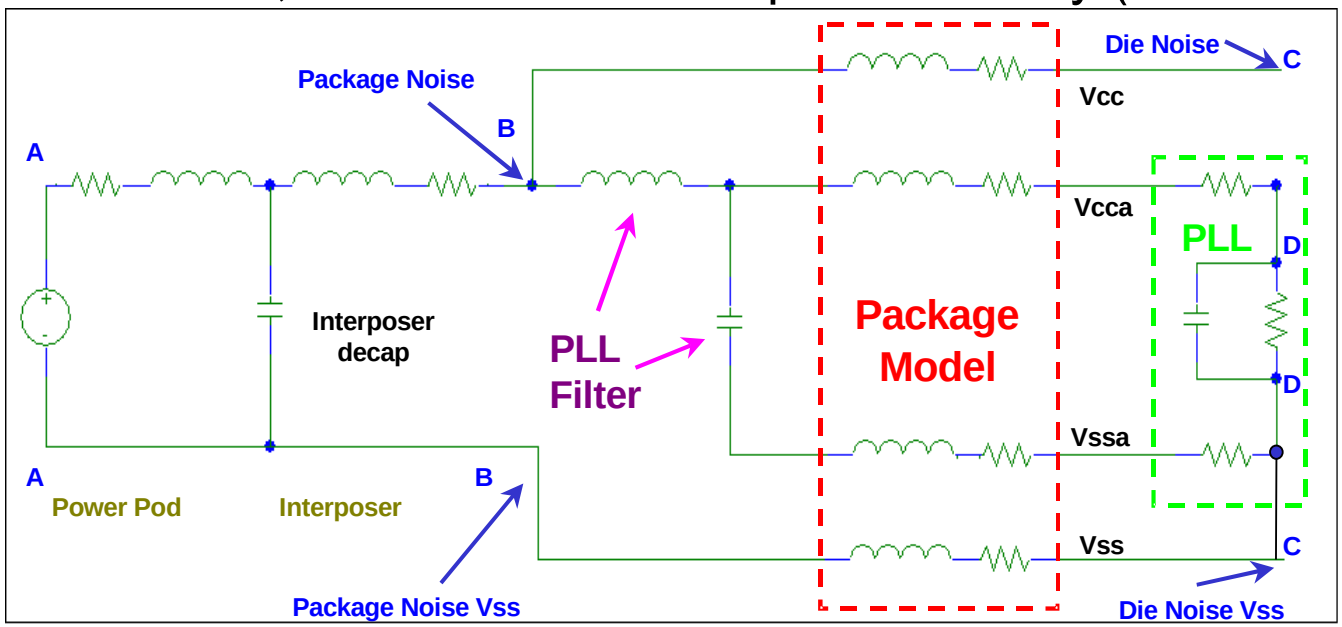
Careful w/ IO Circuit Simulations

- Remember
  - Gnd is an illusion
  - There is not a global reference
- In simulation
  - Chip Vdd/Gnd must be modeled
  - Not equal to board Vdd/Gnd
  - Always measure voltage differen
- Models must reflect true path:
  - Including signal return path
    - In Vdd/Vss network
  - Only way to properly reflect the interaction of Vdd (core supply) and Vtt (IO supply)
  - IO signaling will inject noise into the core Vss (and vice-versa)



Caution with Filtered Supplies

- Certain sensitive circuits need very quiet supplies
  - Examples are PLL's and DLL's
- Desire is to make supplies separate
  - And filter the quiet supply
- This is hard, since Vss often coupled internally (substrate)



References

[1] Kedzierski, et al., paper 10.1, IEDM 2002

[2] Krivokapic, et al., paper 10.7, IEDM 2002

[3] Ng, et al., Table 1, paper 9.6, IEDM 2002

[4] Ishikawa, et al., paper 9.7, IEDM 2002