Power Distribution Networks

Ken Stevens
Design of Power Distribution Networks

- Distribution of power and ground from chip pads to all devices
- Increasingly important design consideration
- Voltage drop due to current drawn by devices
  1. IR drop ($V_{IR}$): voltage drop due to resistance in network
  2. Inductance ($V_{dI/dt}$): voltage drop due to package lead inductance

$$V_{device} = V_{dd} - V_{dI/dt} - V_{IR}$$
Negative Effects of Voltage Drop

1. Switching time of gate increases
   - typical drop of 5-10% – but pattern dependent!!
2. Noise margins reduced
3. Noise injected between devices
4. Metal migration due to excessive currents
5. Resonant noise
Insane Currents

- Current microprocessors:
  - expend up to 120 Watts
  - run at supply of 1.2 V
- Gives 100 Amps!!
- Power scaling makes this worse!
Example Power Supply Noise Waveforms

- Two different Micron packages
- Simple ALU
- 65nm IBM 10sf process
Example Power Supply Noise Waveforms

- Two different Micron packages
- Simple ALU
- 65nm IBM 10sf process
Simplified Power Grid Model

- Power grid modeled as:
  - battery as power source
  - inductive and resistive leads from battery terminals
  - time varying current source for transistors
- Results in power grid fluctuations
- Power grid has resonant frequency:
  \[ f = \frac{1}{2\pi\sqrt{LC}} \]
Simplified Power Grid Model

\[ \text{Graph showing waveforms with labels for } V_d, V_p, \text{ and time axis } t (\text{ns}) \]
Model Order Reduction Accuracy

- Tradeoff in accuracy and run-time
- HSPICE (green) versus Synopsys HSimPlus simulations
  - “two-phase” approach modeling gates as current source
Model Order Reduction Speedup

- Tradeoff in accuracy and run-time
- Test on scalable multiplier array
  - 65nm IBM 10sf process
Model Order Reduction Speedup

- Tradeoff in accuracy and run-time
- Test on scalable multiplier array
  - 65nm IBM 10sf process
Power Grid Failures

- Caused by excessive voltage drop
- Results in circuit failures
  1. performance failures
     - circuit fails at high frequency
  2. functional failures
     - failures are frequency independent
Power Grid Failures

- Three main failure mechanisms
  1. grid induced noise
  2. current density failures
  3. resonance failures
Power Grid Failures

- Noise for 16-stage 16-bit pipelined fibonacci sequence
- HSPICE (yellow) versus HSImPlus
Power Grid Failures

- Noise for 16-stage 16-bit pipelined fibonacci sequence
- Delay variance from simulation with clean power supply
- First droop noise is up to 550 ps!!!
Grid Induced Noise

- Output levels will follow power and ground noise
- Difference in network between transistors can make $V_{gs} > 0$
  - ground can increase in one gate and decrease in another
  - $\Delta$ normally increases with physical distance
  - increases leakage
  - can result in device switching (low $V_t$, domino . . .)
  - decrease in performance for high transitions
- Adds to noise based on coupling, etc.
Current Density Failures

- High currents cause **Electromigration**
  - the redistribution of metal molecules
- Worse in wires where current is unidirectional
- Over time, electromigration causes:
  - increase in resistance
  - open circuits
Resonance Failures

- Resonance can occur in power supply
  \[ f = \frac{1}{2\pi\sqrt{LC}} \]
- Based on values of
  - package inductance
  - on-chip caps (sources, decoupling caps, etc.)
- Failure modes
  - performance degradation
  - power supply fluctuations
    - high voltages: hot carriers
    - low voltages: performance
Power Grid Topologies

- Grids are very topology sensitive
  - resistance, capacitance, current loops, ...

- Topologies are technology dependent
  - number of metal layers, cap properties, supply locations, currents, ...

- Three main grid topologies exist
  1. routed distribution network
  2. power grids
  3. power planes
Routed Power Distribution Networks

- **Typical Application**
  - limited metal layers, power from peripheral rings
  - two adjacent metal layers route power trunks
  - horizontal power, vertical ground
  - intermeshed with other routing signals

- **Advantages**
  - reduced use of metal resources
  - flexibility in metal widths

- **Disadvantages**
  - need inter-trunk straps
  - inconsistent supply
  - normally have power droops in center of blocks
  - require extensive post-layout analysis
Power Distribution Grids

● Typical Application
  ➣ equal width lanes at fixed pitch
  ➣ often several layers, alternating horizontally & vertically
  ➣ supports ball-grid arrays for supplying power supply inside circuitry

● Advantages
  ➣ many redundant paths (failure tolerant, multi-source & direction, . . .)
  ➣ can be used as signal shielding if routed on many layers
  ➣ short current return paths reduce inductance affects

● Disadvantages
  ➣ utilizes a significant portion of metal (25–40%)
  ➣ blocks many routing paths
Power Planes

- Typical Application
  - two full metal layers dedicated to supply
    - one power plane, one ground plane
  - access holes allow vertical plane crossing of signal/supply vias
  - similar to circuit board designs...

- Advantages
  - excellent current distribution
  - lowest power grid failures (noise, migration, etc.)

- Disadvantages
  - expensive use of metal resources
  - difficulty to cross plane with signals/supplies
  - which layer should this be placed on? Why?
Current Loops

- A switching net causes:
  - power to flows from
    - driver source
    - down wire, to destination cap,
    - coupling with signals and power planes in its path
  - this creates a current loop
  - inductance in loop direct function of area: $I = kA$, $k$ determined by magnetic permeability and geometry of loop.
Current Loops

- Unlike R & C, no closed form models exist for on-chip wire inductance
- Extraction difficult, requires return path loop knowledge
- Internally linked to impedance:
  \[ Z = R + j\omega L \]
- Key issues:
  1. near end rise time vs. propagation velocity down wire
  2. attenuation constant
  \[ \frac{Z_0}{2R_{wire}} \]
- Additional loops caused by crowbar (short circuit) current
- Very sensitive to power grid topology. Which is best??
Power Distribution Design

- Similar to most design tasks, early analysis is essential
  - estimate of requirements: current, block placement, ...
  - accuracy will reduce errors and redesign

- Three basic phases
  1. early analysis
  2. floorplan based analysis
  3. layout based validation
Early Power Distribution Analysis

- Set early metal widths and pitches
- Normally assume uniform chip-wide current distribution
- Peak current overestimated by $5-10 \times$ average
- Construct a regular power grid
- Tradeoffs with pump array, power pads
- Example:
  - reduce width and pitch by $2 \times$
  - same metal density
  - for one design, resulted in 40% reduction in voltage drop
Floorplan Based Power Distribution

- Each FUB (FUnctional Block) has power characterization
  - estimated gate count, previous designs, experienced designer
  - PowerMill transistor level simulation
- Placement of FUBs known
- Normally assume uniform FUB current distribution
- Modify grid to account for nonuniform power distribution
Layout Based Power Distribution Analysis

- Estimate magnitude of voltage drop using physical design
  - extract RC design of power delivery network
  - model transistors as time varying current sources
  - estimate inductive effects
- Get time varying voltages & currents at each distribution point
- Final distribution network tuning
Improving the Distribution Network

1. Resistance
   a. widen wires
   b. add wires
   c. add **decoupling capacitance**
      - improves $V_{IR}$ and $V_{dI/dt}$
      - greatly helps inductance
      - serves as local power source
      - reduces supply noise
      - often in series with small resistance to reduce inductance resonance

2. Inductance
Improving the Distribution Network

1. Resistance

2. Inductance
   a. power supply inductance mainly comes from package
   b. add capacitance
   c. add power pads
   d. increase resistance
Improving the Distribution Network

- Adding more power supply taps
  - Simple ALU design, green uses more package power pins
  - IBM 65nm 10sf process, Micron single layer package
Power Grid Resonance

1. RLC values contribute to power supply noise

2. \[ f = \frac{1}{2\pi \sqrt{LC}} \]

3. This frequency has crossed clock frequency (was faster, now slower)

4. BGA technology substantially reduced inductance from pad supply
Power Grid Resonance

Figure 24.9 Voltage plot of a large RLC power grid and associated frequency content.
Power Grid Resonance

- New problems with $f_{ck} > f_I$
- Ringing
  - temporary resonance takes many cycles to settle
  - switching power states identified as problem
  - macro instruction cycles in phase with clock
  - “clock virus”
Power Grid Resonance

- Extremely difficult to increase resonant frequency
- Can *decrease* frequency
- Clever research: *resonant clocking*
  - equate clock and inductive resonance frequencies
  - significant reduction in clocking power
Power Distribution Analysis

- Power distribution analysis is a *global* problem!
  - current at any node affects all others
  - entire grid needs simultaneous analysis
  - to evaluate droop, also need model of transistors...
- Serious demand on run-time and memory
- Can easily take weeks or months of run time
Power Distribution Analysis

- Dual simulation approach
  1. nonlinear simulation
     - simulate transistors using clean power and ground
     - record current and times
  2. simulation of linear devices (R,C,L)
     - add in currents from nonlinear simulation
Power Distribution Analysis

- Dual simulation is optimistic simulation
  - nonlinear simulated with full rail
  - therefore, overestimates current
  - $\approx$ linear overestimation based on voltage droop

- Can iterate
Linear Solvers

- Multi-layer grid, requiring differential equation solver
- Interesting algorithmic problems of high value
  - modified nodal analysis using backward Euler techniques
  - matrix factorization using direct (Cholesky factorization) and iterative (conjugate gradient) solvers
  - solve AC (direct usually better) or DC (iterative) analysis.
- Best approaches determined by
  - size, matrix sparsity, matrix structure, grid model, AC or DC analysis.
Non-Linear Simulation

- Generate worst realistic voltage drop at each grid location
- Dependent on vectors
- SPICE or SPICE-like simulation engines

Vector generation:
1. user generated hot loops
   - hand design, labor intensive, error prone
2. automated peak current vector generation
   - limited scale, poor physical block instancing information
3. static current generation algorithms
   - ignores logic correlations, uses timing windows
Power Grid Models

- Accuracy of results dependent on model
- Models key to time/accuracy/physical representation
- Resistance model:
  - well characterized by process:
    - pad connections, vias, contacts. metal sheet resistances (Ω per □)
  - $R_{wire} = \frac{\rho L}{W}$, $\rho$ is sheet resistance
Capacitance Models

- Significant caps to power supply include:
  - capacitance between supply wires and other wires and substrate
  - parasitic caps of transistor sources
  - parasitics of wells
  - decoupling capacitors

- Commonly extract these values using 2 or 3-dimensional models

- Normally simplified by “lumping” segments in final models
Statistical Signal Wire Model

- Switching of signals complicates model
- Related to MCF (maximum coupling factor)
- One model given in book:
  - use average of nets that switch in cycle $P_{active}$
  - assume equal high and low switching probability, resulting in canceling
  - assume half of the signals are high, half low voltage, so half will follow rail
  - this results in the following statistical wire parasitic model:
    \[ C_{\text{eff}} = \frac{1}{2} C_{\text{coupling}} (1 - P_{active}) \]
Statistical Coupling Model

- The caps in a gate can also be modeled statistically:
- Ignore $C_{sb}$ if we don’t body bias
- Active transistors modeled by nonlinear current sources
- For inactive node $(1 - P_{active})$, decoupling cap is sum of effective high and low capacitances weighed by the probability of being in each state:
  \[ C_{eff} = \frac{1}{2} (1 - P_{active}) (C_{gd} + C_{gs}) \]
- Well cap can be modeled as $R_{eff} C_{eff}$ where $C_{eff}$ is well junction cap and the resistance depends on the placement of well taps.
Decoupling Caps

- Use gate as capacitor
- Gate oxide breakdown controlled with fuse or transistor in series.
- Model series resistance
  - increase to dampen oscillations
  - decrease to improve $dI/dt$ noise reduction
- Other interesting approaches: fractal designs...
Package Inductance Models

- Include series inductance and resistance of package
- Package decoupling caps not effective above 500MHz (need on-chip decap)
- Large model for bump array power supplies
- Model Vdd and Gnd simultaneously to get inductive effects
Power Grid Model for PPC 750

- Using R (a), RC (b), and RLC (c) models
Power Distribution Network Review

1. Two major sources of voltage supply droop
2. Negative effects of voltage supply droop
3. Characteristics of voltage droop failures
4. Three main failure mechanisms
5. Electromigration
6. Power supply resonant frequency $f = \frac{1}{2\pi \sqrt{LC}}$
7. How to change resonant frequency
Power Distribution Network Review

1. Failure modes due to power supply resonantion
2. Three main power grid design styles, advantages and disadvantages
3. Current loops
4. Best power distribution design methodology
5. Two design parameters for improving distribution networks
6. Decoupling capacitors, and their tradeoffs
Power Distribution Network Review

1. Design methods for changing inductance
2. Clock resonance problems with high frequency clocks
3. Challenges in analyzing power distribution networks
4. Two-level approach for grid simulation
5. Methods to reduce complexity
6. Decoupling cap design