Technology Trends in Power-Grid-Induced Noise

Sani R. Nassif  
IBM Research – Austin  
nassif@us.ibm.com

Onsi Fakhouri  
MIT  
onsi@mit.edu
ITRS Power Fashion Statement

Farid Najm (from *TRS data)
Motivation

- Power is increasing (hot plates, nuclear reactors etc...).
- $V_{DD}$ is decreasing ($V_{TH}$ decreasing slower to manage leakage).
- Frequency is increasing.
- Dynamic and static $I_{DD}$ are increasing (electromigration!).

- IR and Ldi/dt noise becoming a larger part of the total noise budget.
- Impact of $V_{DD}$ variation on delay is increasing.
  - (Because of reduced overdrive $V_{DD} - V_{TH}$)
- Understanding the origins and trends of supply induced noise becoming critical.
Leakage Current “Predictions”

Tak Ning (from *TRS data)
Outline

- Expressions for power grid induced noise.
- Technology trends.
- Design realities and trends...
- Power Grid Planning.
- Power Grid Planning Examples.
- Open issues and low-hanging fruit.
Canonical Power Grid Circuit

- Grid is predominantly **resistive**.
- Package is predominantly **inductive**.
- Load is current.
- Other circuits ~ lossy decoupling capacitance.
Grid Capacitance

- Capacitance ~ 1/distance.

- Distance scale for power grid is in the range of 10\(\mu\)m.

- Distance scale for device capacitance is in the range of \(T_{\text{OX}}\) ~ 20nm.

- Capacitance “density” of devices makes grid capacitance unimportant.
Grid Inductance

- Rectangular Conductor...
- Worst case for grid wire.
- \( L_{(pH)} \approx 0.2 \ l \ \ln(2l/(w+t) + 0.5) \)
- Package parasitics much greater.

Range of Package

![Graph showing the relationship between length (microns) and inductance (pH).]
Grid Resistance

- Same Conductor...
- \( R \sim 0.1 \, \text{l/(w t)} \)
- Package parasitics **much smaller**.
Noise Model

- **Current modeled as:**
  - \( I = 0 \quad t < 0 \)
  - \( I = \mu t \quad t < t_p \)
  - \( I = \mu(2t_p-t) \quad t < 2 t_p \)
  - \( I = 0 \quad t > 2 t_p \)

- **Ignoring L, maximum noise is:**
  \[
  V_{\text{max}} = \mu t_p R_g - \mu R_g^2 C_d (1 - e^{-t_p/\tau})
  \]
  \[\tau = (R_g + R_d) C_d\]
  \(\approx \mu t_p R_g / (R_g + R_d)\)
  (for large \( C_d \))
With package, maximum noise becomes:

\[ V_{\text{max}} \approx \mu t_p R_g + \mu L - \mu R_g^2 C_d (1 - e^{-t_p/\tau}) \]

Accurate expression:

\[ V_{\text{max}} = \mu t_p R_g + \mu L - \mu R_g^2 C_d + \psi_1 + \psi_2 \]

\[ e_1 = \exp - (\tau + \beta) t_p / 2C_d L \quad e_2 = \exp - (\tau - \beta) t_p / 2C_d L \]

\[ \beta = (\tau^2 - 4LC_d)^{\frac{1}{2}} \]

\[ \psi_1 = (e_1 + e_2) \mu (L - C_d R_g^2) / 2 \]

\[ \psi_2 = (e_1 - e_2) \mu C_d (\tau R_g^2 - L(3R_g - R_d)) / 2\beta \]
Quality of Approximation

Good accuracy where needed!
Outline

- Expressions for power grid induced noise.
- Technology trends.

- Design realities and trends...
- Power Grid Planning.
- Power Grid Planning Examples.

- Open issues and low-hanging fruit.
Technology Variables

- We need to find trends in the parameters of our canonical model.
- Roadmaps provide insight into $V_{DD}$, Area, Power, Frequency etc...
## Technology Parameters

<table>
<thead>
<tr>
<th>Year</th>
<th>L(_{\text{EFF}}) (nm)</th>
<th>F (MHz)</th>
<th>V(_{\text{DD}}) (Volts)</th>
<th>Area ((\text{}\text{\text{2mm}}))</th>
<th>Power (Watts)</th>
<th>Power Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>140</td>
<td>1200</td>
<td>1.8</td>
<td>450</td>
<td>90</td>
<td>0.2</td>
</tr>
<tr>
<td>2000</td>
<td>120</td>
<td>1321</td>
<td>1.8</td>
<td>450</td>
<td>100</td>
<td>0.22</td>
</tr>
<tr>
<td>2001</td>
<td>100</td>
<td>1454</td>
<td>1.5</td>
<td>450</td>
<td>115</td>
<td>0.26</td>
</tr>
<tr>
<td>2002</td>
<td>85</td>
<td>1600</td>
<td>1.5</td>
<td>509</td>
<td>130</td>
<td>0.26</td>
</tr>
<tr>
<td>2003</td>
<td>80</td>
<td>1724</td>
<td>1.5</td>
<td>567</td>
<td>140</td>
<td>0.25</td>
</tr>
<tr>
<td>2004</td>
<td>70</td>
<td>1857</td>
<td>1.2</td>
<td>595</td>
<td>150</td>
<td>0.25</td>
</tr>
<tr>
<td>2005</td>
<td>65</td>
<td>2000</td>
<td>1.2</td>
<td>622</td>
<td>160</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Dependencies

- Time $t_p \sim F^{-1}$
- Power density $P_\square \sim V_{DD} \mu \ t_p \rightarrow \mu \sim F \frac{P_\square}{V_{DD}}$
- $C_d \sim C_{OX} \sim 1/T_{OX}$
- $R_d$ and $R_g$ are $\sim$ constant
  - But proper power grid planning can make a difference here!
- $L$ is $\sim$ constant
  - Package learning curve is much shallower than technology learning curve!
Noise Trends

Based on conservative ITRS trends

\[ V_{DD} \approx 0.6X \]
\[ t_p \approx 0.6X \]
\[ \mu \approx 3.3X \]
\[ C_d \approx 2X \]

\[ V_{max} \approx \mu t_p R_g + \mu L - \mu R_g^2 C_d (1 - e^{-t_p/\tau}) \]

\[ \approx 2X \quad \approx 3X \quad \approx \text{Same} \]

\[ \text{Plus} \approx 1.7X \text{ due to reduction in } V_{DD} \]
Outline

Expressions for power grid induced noise.
Technology trends.

Design realities and trends...
Power Grid Planning.
Power Grid Planning Examples.

Open issues and low-hanging fruit.
Power Grid Design Trends

- Number of levels of metal is increasing.
  - More degrees of freedom for tradeoff between interconnect and power.
  - More effort in grid design.
- Cu and advanced CMP processes place more design restrictions on wires.
  - Example: maximum width, metal density, oxide density within metal area, etc...
- Number of package power pins for high power chips increasing (fixed $I_{\text{max}}$ per pin).
Package Choices

Area Array (C4)
- Power distributed across all the chip area.

Wirebond (periphery)
- Power brought in from edge of chip.
**SOC and IP Constraints**

- Hard IP places constraints and creates discontinuities in grid design.
- Often dealt with using “rings” (area hit).
Power Grid Design Issues

- Power Grid impacts implementation of every component at the PD level.
- Placement of power-hungry devices (I/O buffers, clocks, etc...).
- Placement and allocation of decoupling capacitors to minimize noise.
- “Interface” between incompatible power distributions costly in routing resources.

- It is not unthinkable to use 15 to 20% of wiring resources for power distribution.
Buffer Placement Algorithm

- ICECS ‘00 paper (J. Kozhaya et. al.)
- A greedy heuristic technique.
- Idea: Use sensitivity information to place I/O buffers one at a time while satisfying drop thresholds.
- The $A^{-1}$ (system matrix) provides sensitivity of voltage drops to placement of I/O buffers.
  - I/O buffers only appear in the RHS of the system of linear equations!
1. Sort I/O buffers and initialize drop slacks.
2. For buffer $B_k$, compute upper bounds on the allowable current at every node $n_i$ which is a potential placement site.
3. Assign buffer $B_k$ to node $n_m$ where $n_m$ is the node with the largest upper bound.
4. Update the drop slack at all nodes:
   - $s(j) = s(j) - a_{jm}^{-1} I_k$, $\forall j$
5. If $s(j) < 0$, report a violation at node $n_j$.
6. Continue at step 2 with the next buffer.
Results

<table>
<thead>
<tr>
<th>Design</th>
<th># Buffers</th>
<th># Nodes</th>
<th>Violations</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 (0.18µ)</td>
<td>616</td>
<td>4602</td>
<td>0</td>
<td>3.79</td>
</tr>
<tr>
<td>C2 (0.13µ)</td>
<td>588</td>
<td>3325</td>
<td>0</td>
<td>3.04</td>
</tr>
</tbody>
</table>

- Technique finds a *feasible* placement.
- CPU time is fast enough for iteration.
- Results were verified using detailed simulation.
Is This an Easy Problem?

Results of placing the I/Os randomly with 1.0% drop thresholds:
Decoupling Capacitance Sizing

Upcoming DAC’02 paper (H. Su et al.)

Before optimization

After optimization
Impact of Decap Sizing

Decap optimization allows ~10% more circuit density!
Outline

- Expressions for power grid induced noise.
- Technology trends.

- Design realities and trends...
- Power Grid Planning.
- Power Grid Planning Examples.

- Open issues and low-hanging fruit.
Power Grid Planning

- Power grid is usually designed **BEFORE** detailed implementation has started.

- Predefined “Image” for ASIC or SOC implementations.

- “Tile” based image for custom ICs.

- Grid is defined at a time when the spatial information about power requirements is approximate, therefore rampant overdesign!
IBM Power Grid Planner

Spreadsheet-like interface to define overall power grid.
IBM Power Grid Planner

Lots of Visualization and Analysis...
IBM Power Grid Planner

- Usually used to explore design options very early in the design cycle.
- Tool needs to be very fast (interactive).

Typical questions:
- Can a grid with X% density handle P watts per square mm?
- How much decoupling capacitance does an I/O buffer need? How close does it need to be?
- How much do I gain by introducing skew?
Outline

- Expressions for power grid induced noise.
- Technology trends.
- Design realities and trends...
- Power Grid Planning.
- Power Grid Planning Examples.
- Open issues and low-hanging fruit.
Planning Examples

Question:
- What is the impact of wiring resources on a per layer basis?

Methodology:
- Perform a full factorial experiment varying wiring density on each level from 5% to 20% and measure grid performance.
- Build a statistical model of grid performance vs. layer densities.
Experiment

- 7 level metal process.
- Top level fixed to interface to package C4’s.

- Density $\sim$ width/pitch.
- Pitch goes up by 2X every 2 layers $(1,1,2,2,4,4)$
- $4^6 = 4096$ simulation $\sim$ 10 hours CPU time.

- Measure VDD and GND net statistics.
Example of Results

![Graph showing Mean VDD drop and Density on Layer 6]
## Analysis of Results

Linear regression of noise vs. layer densities.

<table>
<thead>
<tr>
<th>name</th>
<th>vddmax</th>
<th>vddmean</th>
<th>gndmax</th>
<th>gndmean</th>
</tr>
</thead>
<tbody>
<tr>
<td>rho</td>
<td>0.948</td>
<td>0.941</td>
<td>0.939</td>
<td>0.935</td>
</tr>
<tr>
<td>range</td>
<td>0.0527,0.1804</td>
<td>0.0364,0.1142</td>
<td>0.0360,0.1296</td>
<td>0.0182,0.0585</td>
</tr>
<tr>
<td>K</td>
<td>0.18164</td>
<td>0.12115</td>
<td>0.12010</td>
<td>0.05549</td>
</tr>
<tr>
<td>d0</td>
<td>-0.02428</td>
<td>-0.00001</td>
<td>-0.02616</td>
<td>-0.00520</td>
</tr>
<tr>
<td>d1</td>
<td>-0.10201</td>
<td>-0.03502</td>
<td>-0.13665</td>
<td>-0.06095</td>
</tr>
<tr>
<td>d2</td>
<td>-0.02937</td>
<td>-0.00537</td>
<td>-0.02764</td>
<td>-0.00739</td>
</tr>
<tr>
<td>d3</td>
<td>-0.12701</td>
<td>-0.06568</td>
<td>-0.13211</td>
<td>-0.06287</td>
</tr>
<tr>
<td>d4</td>
<td>-0.05253</td>
<td>-0.03394</td>
<td>-0.02857</td>
<td>-0.00947</td>
</tr>
<tr>
<td>d5</td>
<td>-0.39699</td>
<td>-0.33490</td>
<td>-0.13706</td>
<td>-0.06973</td>
</tr>
</tbody>
</table>

Directional dependence! (anisotropy)
Outline

Expressions for power grid induced noise.
Technology trends.

Design realities and trends...
Power Grid Planning.
Power Grid Planning Examples.

Open issues and low-hanging fruit.
Open Issues

- Coupling of power and timing results.
  - Some early results, but nothing real yet!
- Fast modeling and prediction of chip/package resonance.
  - Approximations OK, but better numerical analysis can make results more accurate.
- Vector-less Chip-level power estimation.
  - Most design flows are not yet focused on power. Need a method to jumpstart power analysis.
- Coupling of power and thermal results + impact on reliability.
  - Same math, same inputs...
Low Hanging Fruit

- Improved modeling of load currents and decoupling capacitance.
  - Holistic rationalization of accuracy requirements.
- Integration of placement aspects of PD with power grid analysis.
  - Moving loads around can be done very efficiently (new RHS and forward/backward solver of existing LU factors).
- Use of sparse inductance formulations to speed up chip/package analysis.
  - Reuse of existing simulation infrastructure.