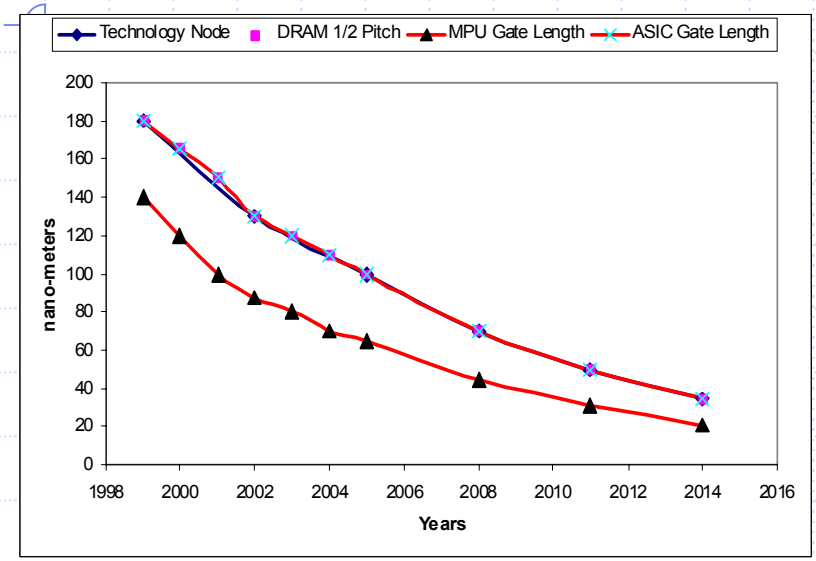
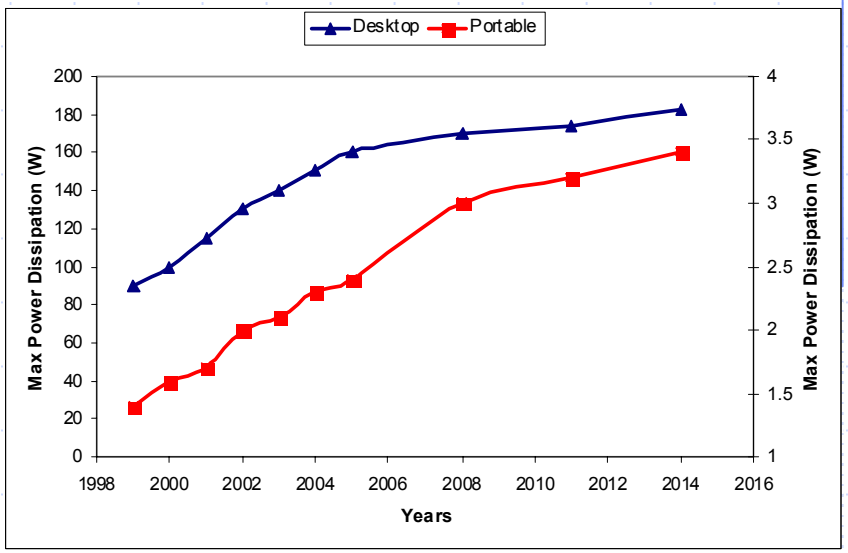
Technology Trends in Power-Grid-Induced Noise

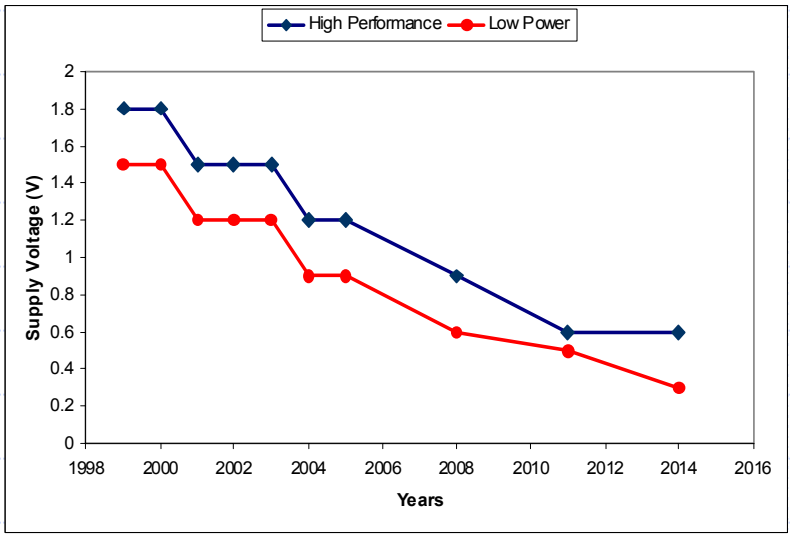
Sani R. Nassif
IBM Research – Austin
nassif@us.ibm.com

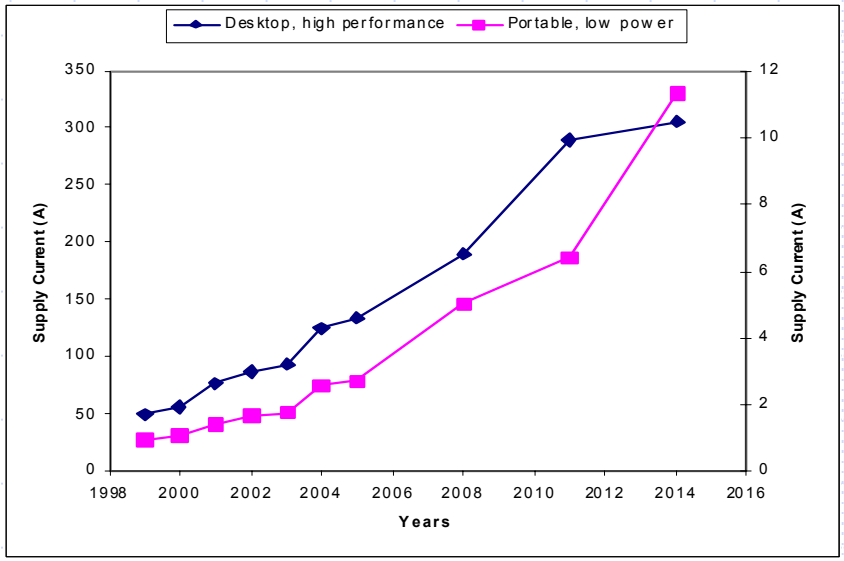
Onsi Fakhouri
MIT
onsi@mit.edu

ITRS Power Fashion Statement









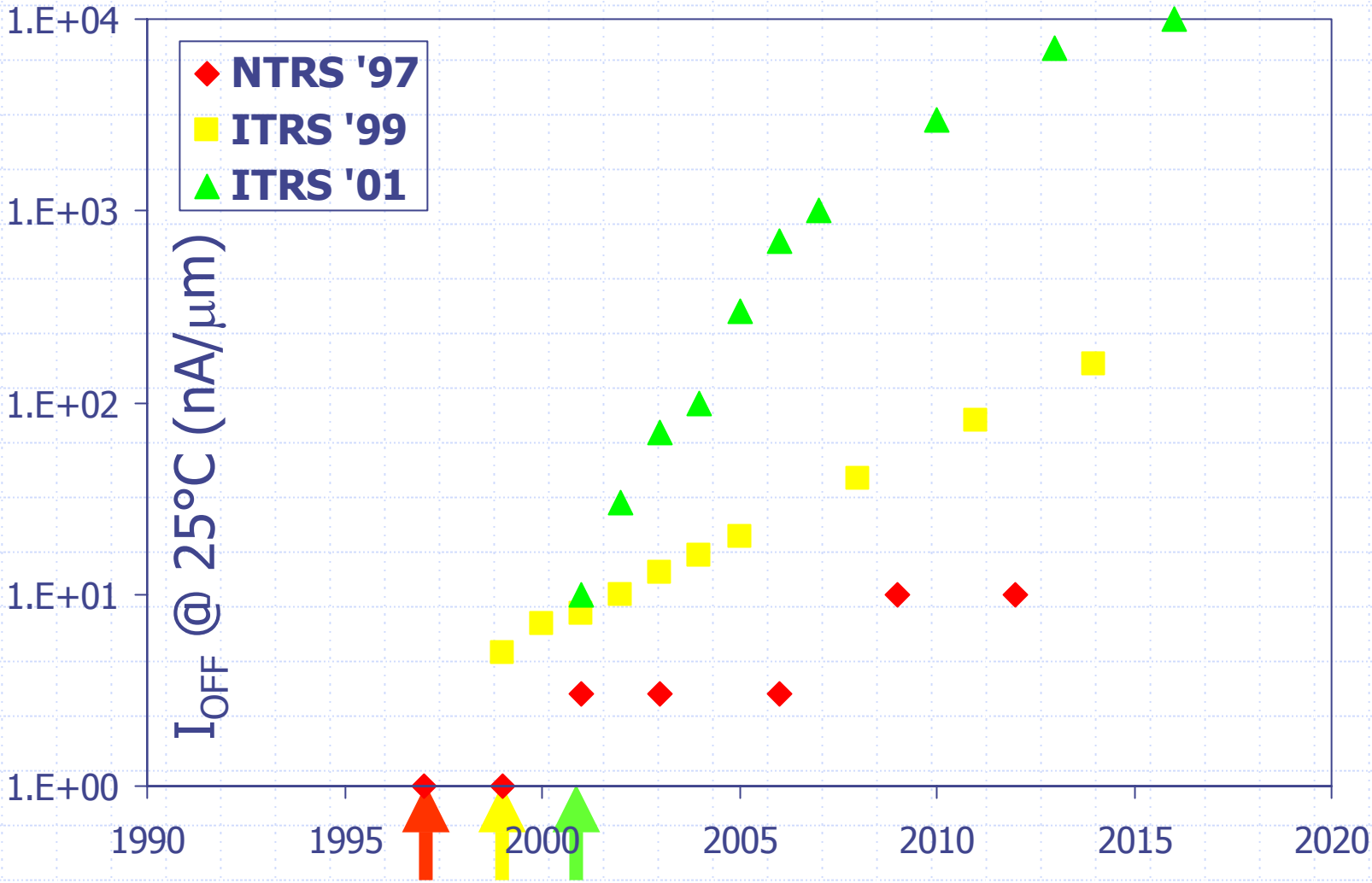
Farid Najm (from *TRS data)

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

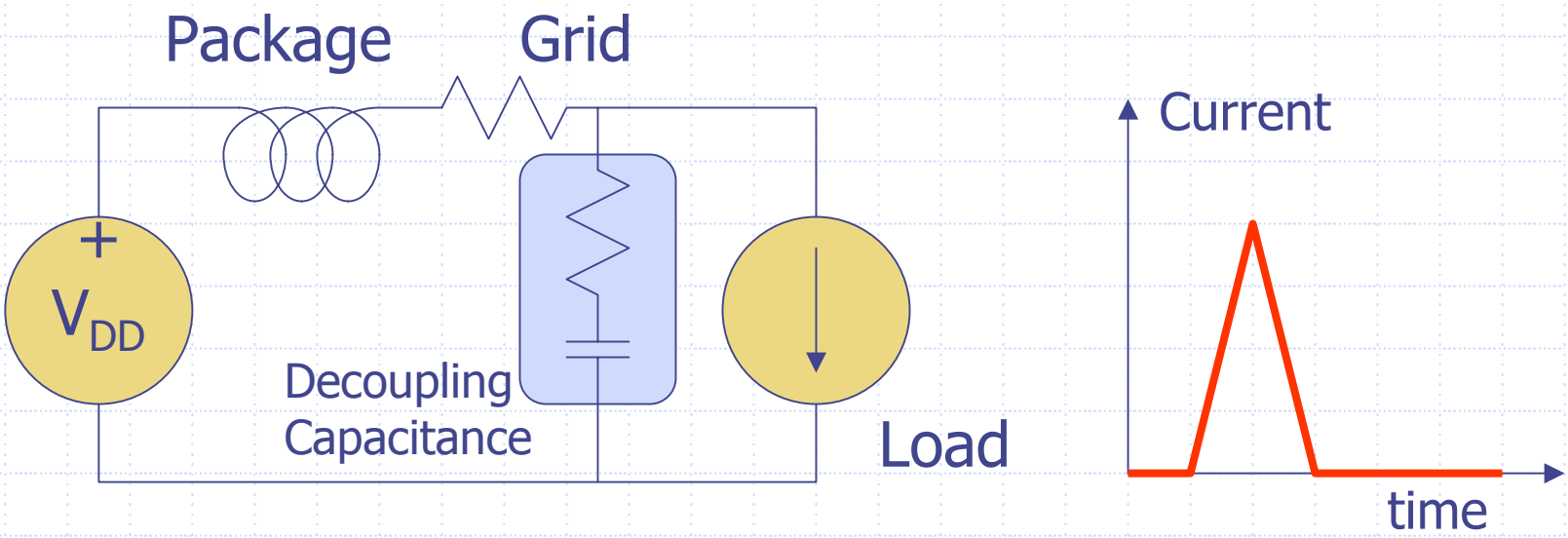
- Expressions for power grid induced noise.
- Technology trends.

- 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.

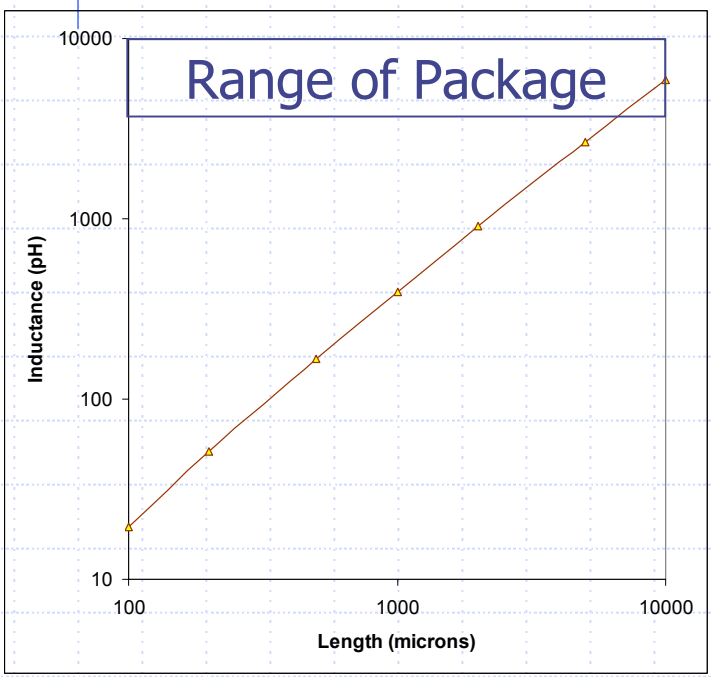


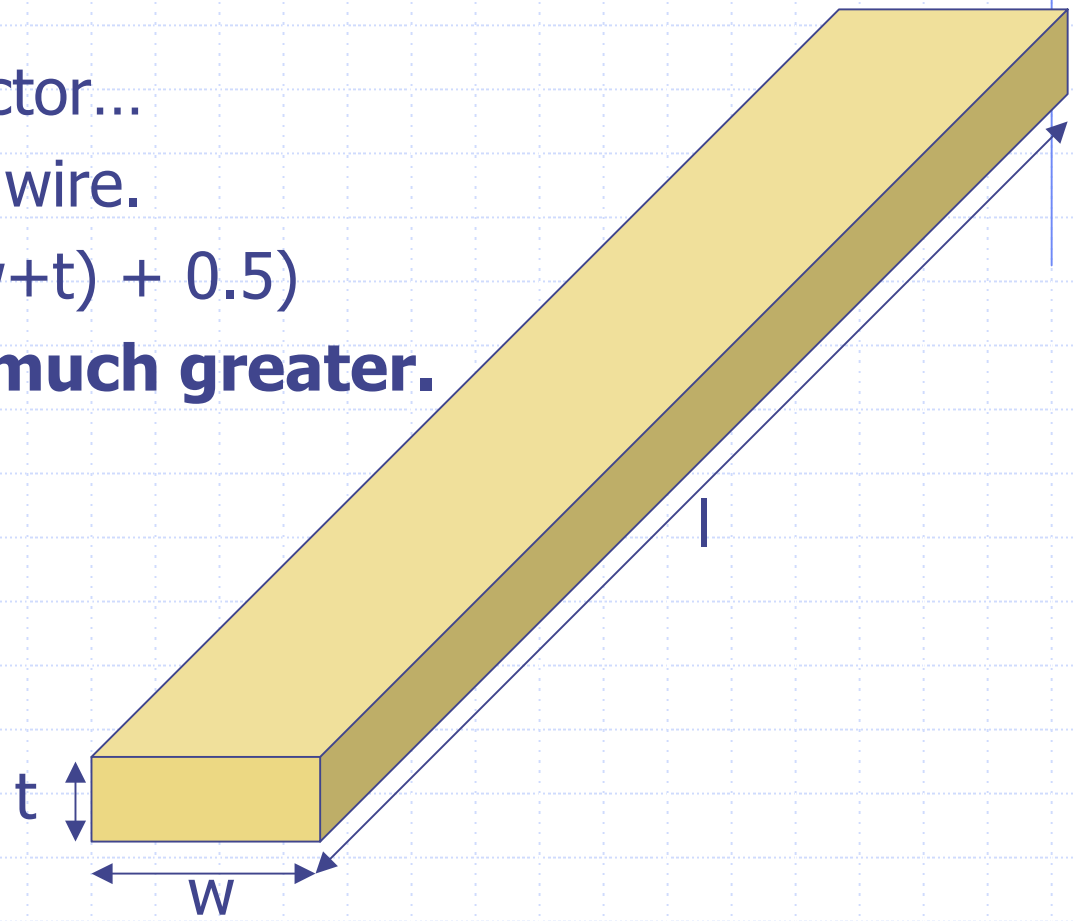
Grid Capacitance

- Capacitance ~ 1/distance.
- Distance scale for power grid is in the range of 10μm.
- Distance scale for device capacitance is in the range of T_{OX} ~ 20**n**m.
- Capacitance "density" of devices makes grid capacitance unimportant.

Grid Inductance

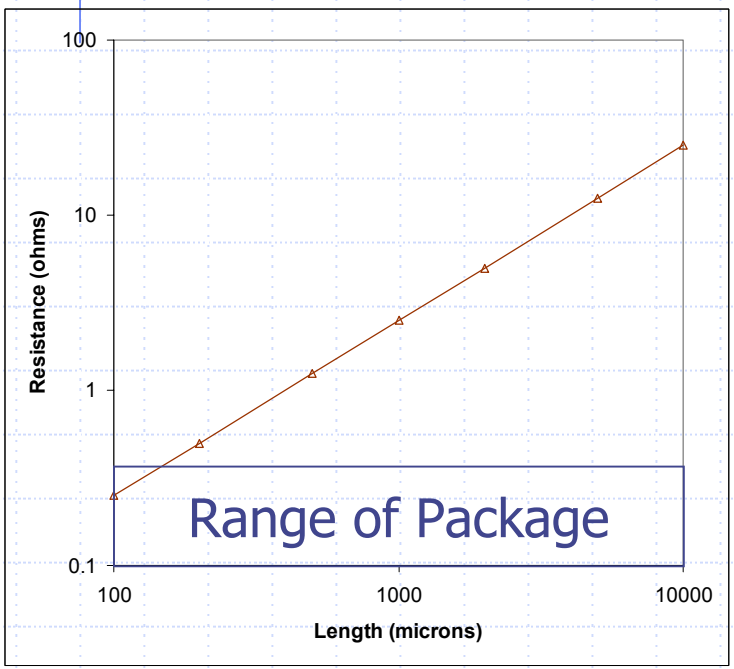
- Rectangular Conductor…
- Worst case for grid wire.
- $L_{(pH)} \sim 0.2\ l\ \ln(2l/(w+t) + 0.5)$
- Package parasitics **much greater.**

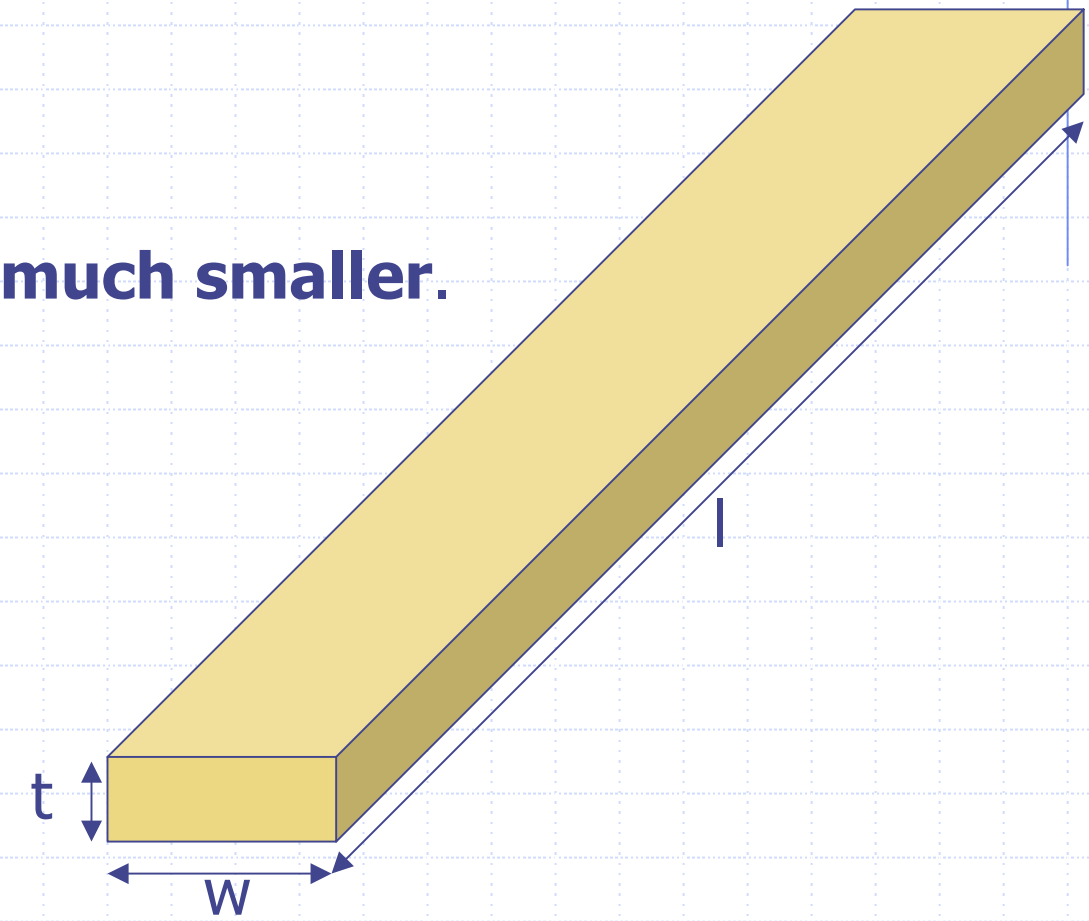




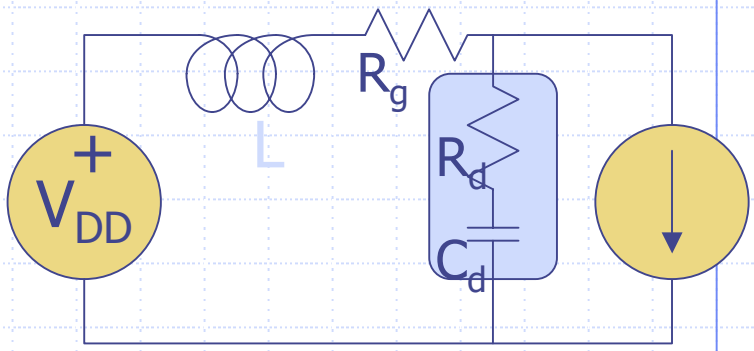
Grid Resistance

- Same Conductor…
- R ~ 0.1 l/(w t)
- Package parasitics **much smaller**.



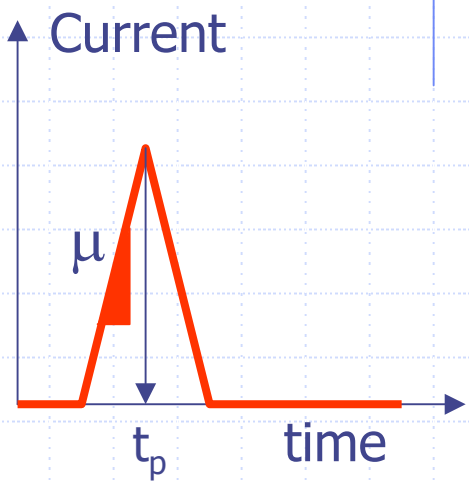


Noise Model



- Current modeled as:

$I = 0$	$t < 0$
$I = \mu t$	$t < t_p$
$I = \mu(2t_p - t)$	$t < 2\,t_p$
$I = 0$	$t > 2\,t_p$



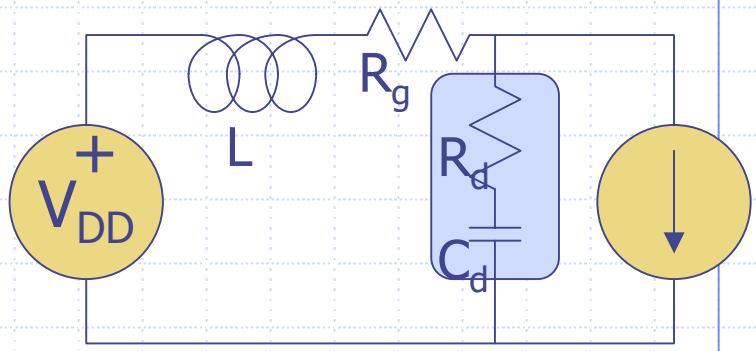
- Ignoring L, maximum noise is:

$$V_{max} = \underbrace{\mu\, t_p\, R_g}_{DC} - \underbrace{\mu\, R^2_g\, C_d\,(1 - e^{-t_p/\tau})}_{Decap}$$

$$\approx \mu\, t_p\, R_g / (R_g + R_d)$$
(for large C_d)

$$\tau = (R_g + R_d)\, C_d$$

Noise Model + L



- With package, maximum noise becomes:

$$V_{max} \approx \underbrace{\mu\, t_p\, R_g}_{DC} + \underbrace{\mu L}_{Package} - \underbrace{\mu\, R^2_g\, C_d\, (1 - e^{-t_p/\tau})}_{Decap}$$

- Accurate expression:

$$V_{max} = \mu\, t_p\, R_g + \mu L - \mu\, R^2_g\, C_d + \Psi_1 + \Psi_2$$

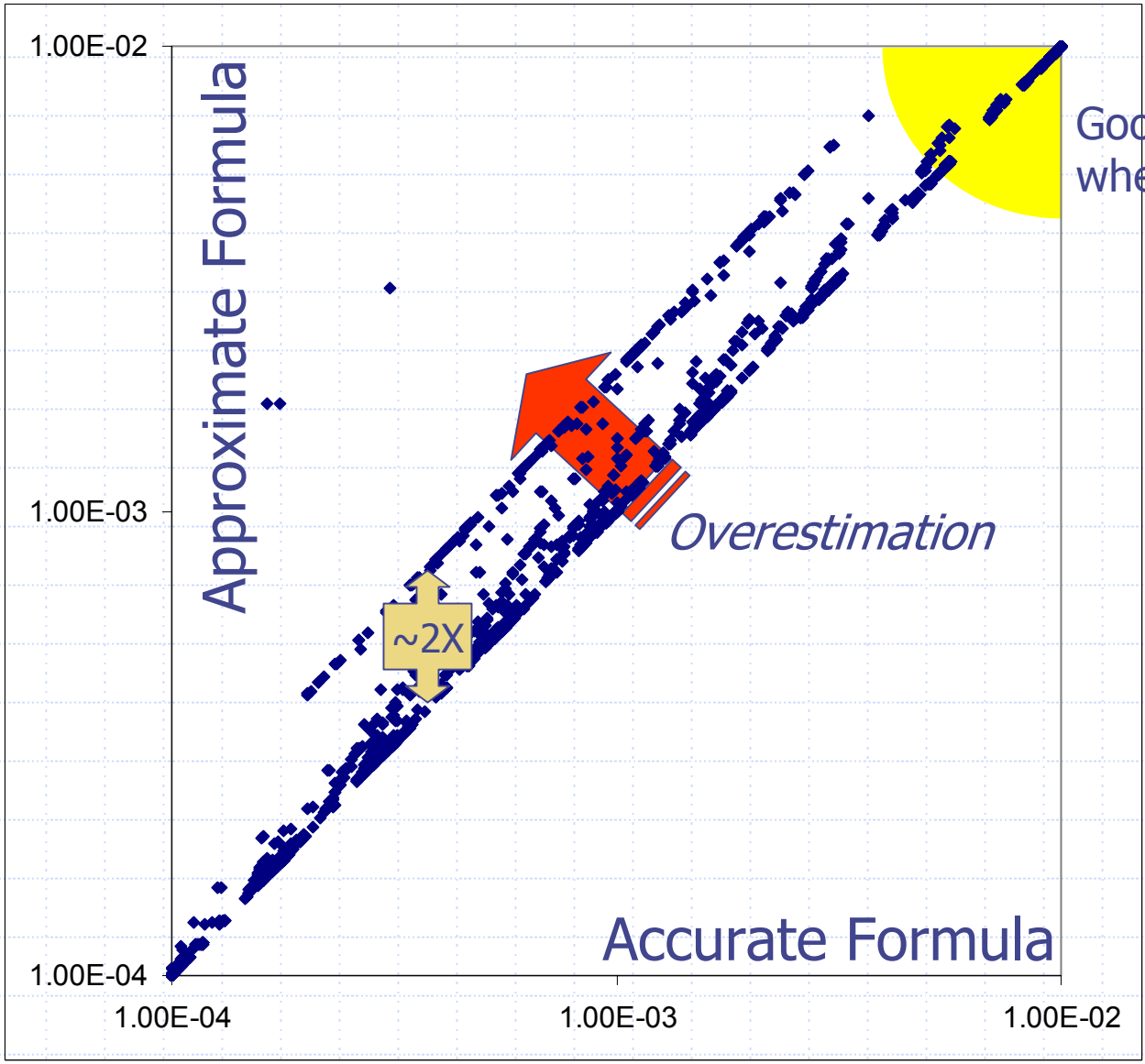
$$e_1 = \exp -(\tau+\beta)t_p/2C_dL \qquad e_2 = \exp -(\tau-\beta)t_p/2C_dL$$

$$\beta = (\tau^2 - 4LC_d)^{1/2}$$

$$\Psi_1 = (e_1 + e_2)\, \mu\, (L - C_d\, R_g^2) / 2$$

$$\Psi_2 = (e_1 - e_2)\, \mu\, C_d\, (\tau\, R_g^2 - L(3R_g - R_d)) / 2\beta$$

Quality of Approximation



Outline

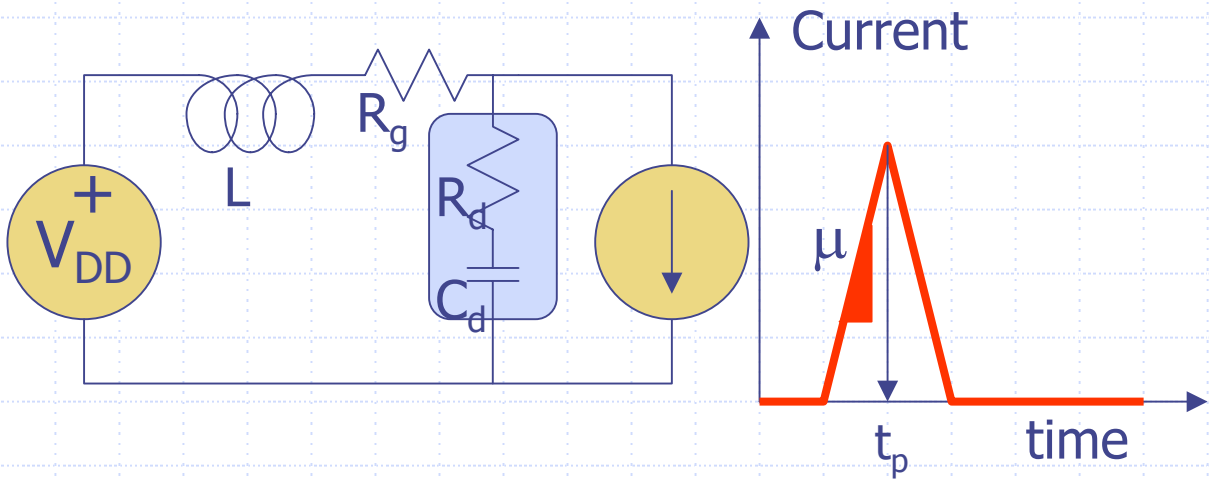
- Expressions for power grid induced noise.
- Technology trends.

- 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…



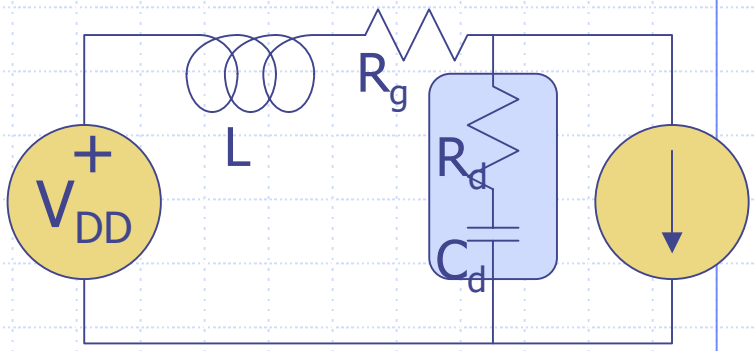
Technology Parameters

Year	L_{EFF} (nm)	F (MHz)	V_{DD} (Volts)	Area (^{2}mm)	Power (Watts)	Power Density
1999	140	1200	1.8	450	90	0.2
2000	120	1321	1.8	450	100	0.22
2001	100	1454	1.5	450	115	0.26
2002	85	1600	1.5	509	130	0.26
2003	80	1724	1.5	567	140	0.25
2004	70	1857	1.2	595	150	0.25
2005	65	2000	1.2	622	160	0.26

Dependencies

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

Noise Trends



V_{DD} 0.6X

t_p 0.6X

μ 3.3X

C_d 2X

Based on conservative ITRS trends

~2X ~3X

$$V_{max} \approx \underbrace{\mu\, t_p\, R_g}_{DC} + \underbrace{\mu L}_{Package} - \underbrace{\mu\, R^2_g\, C_d\,(1 - e^{-t_p/\tau})}_{Decap} \quad \sim Same$$

Plus ~ 1.7X due to reduction in V_{DD}

Outline

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

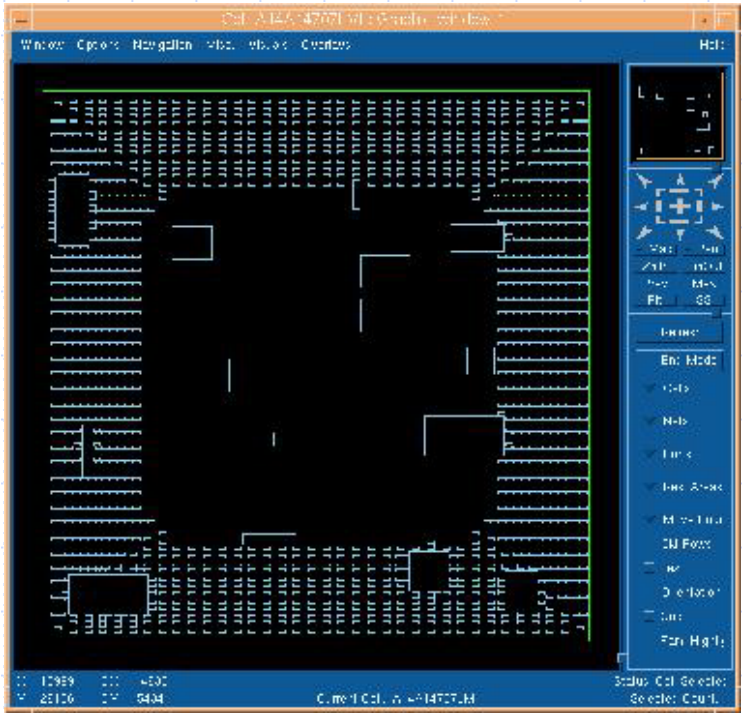
- Design realities and trends…
- Power Grid Planning.
- Power Grid Planning Examples.

- Open issues and low-hanging fruit.

Power Grid Design Trends

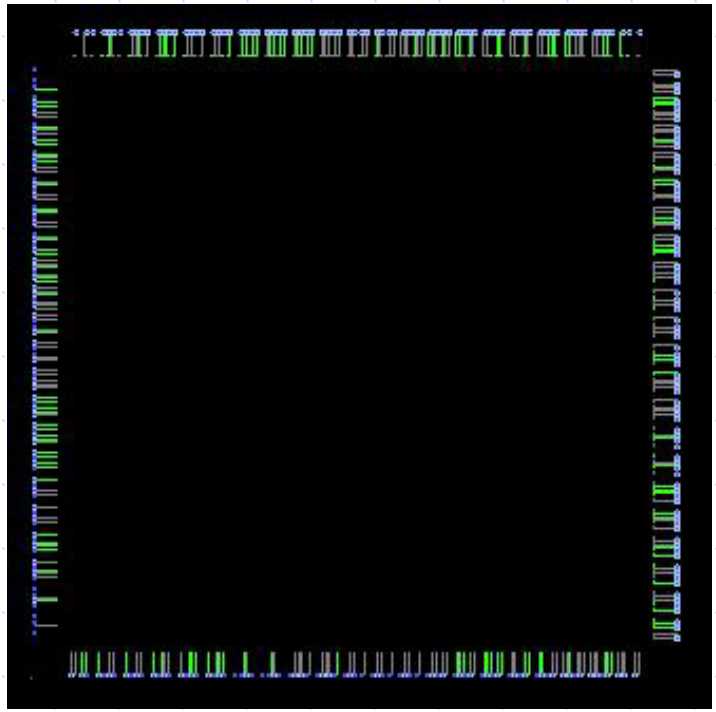
- Number of levels of metal is increasing.
  - More degrees of freedom for tradeoff between interconnect and power.
  - More effort in grid design.
- Cu and advanced CMP processes place more design restrictions on wires.
  - Example: maximum width, metal density, oxide density within metal area, etc…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

Package Choices



Area Array (C4)

- Power distributed across all the chip area.

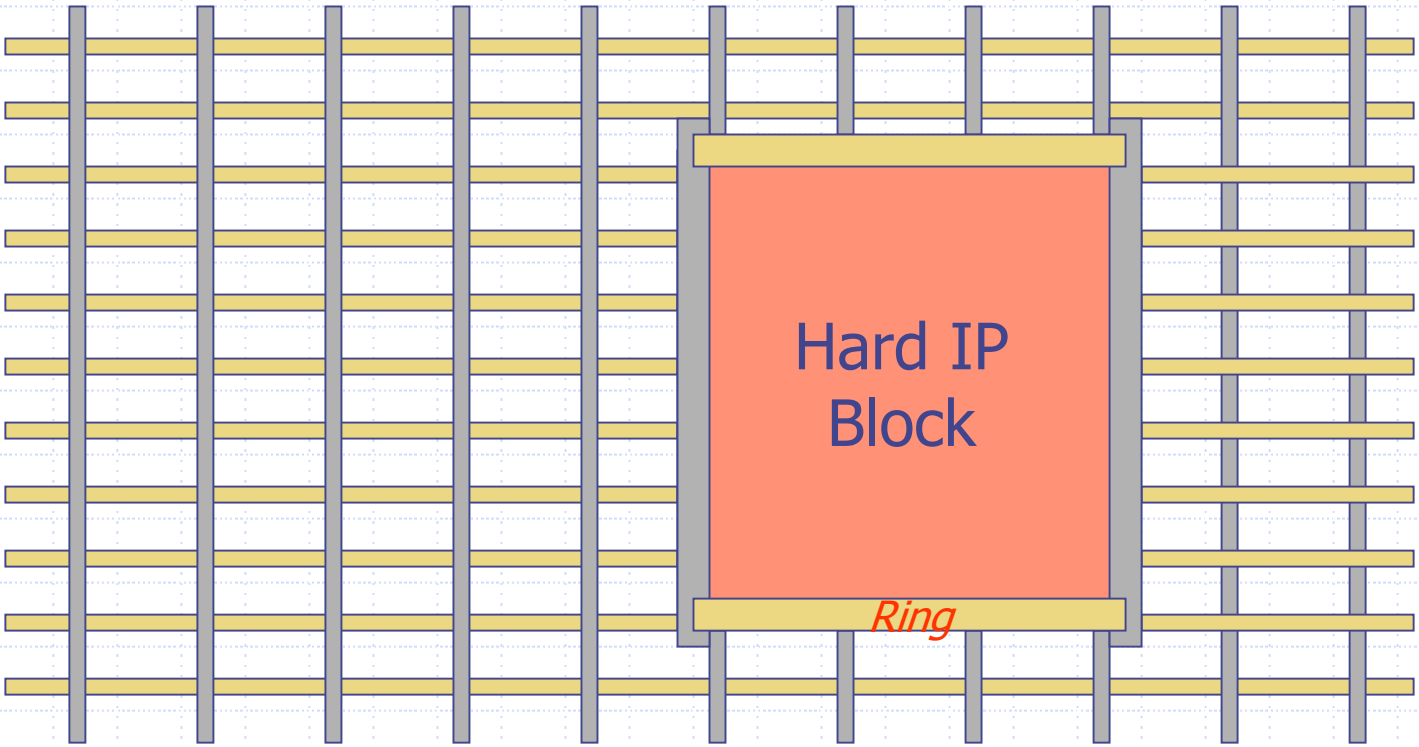


Wirebond (periphery)

- 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!

Algorithm Description

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.

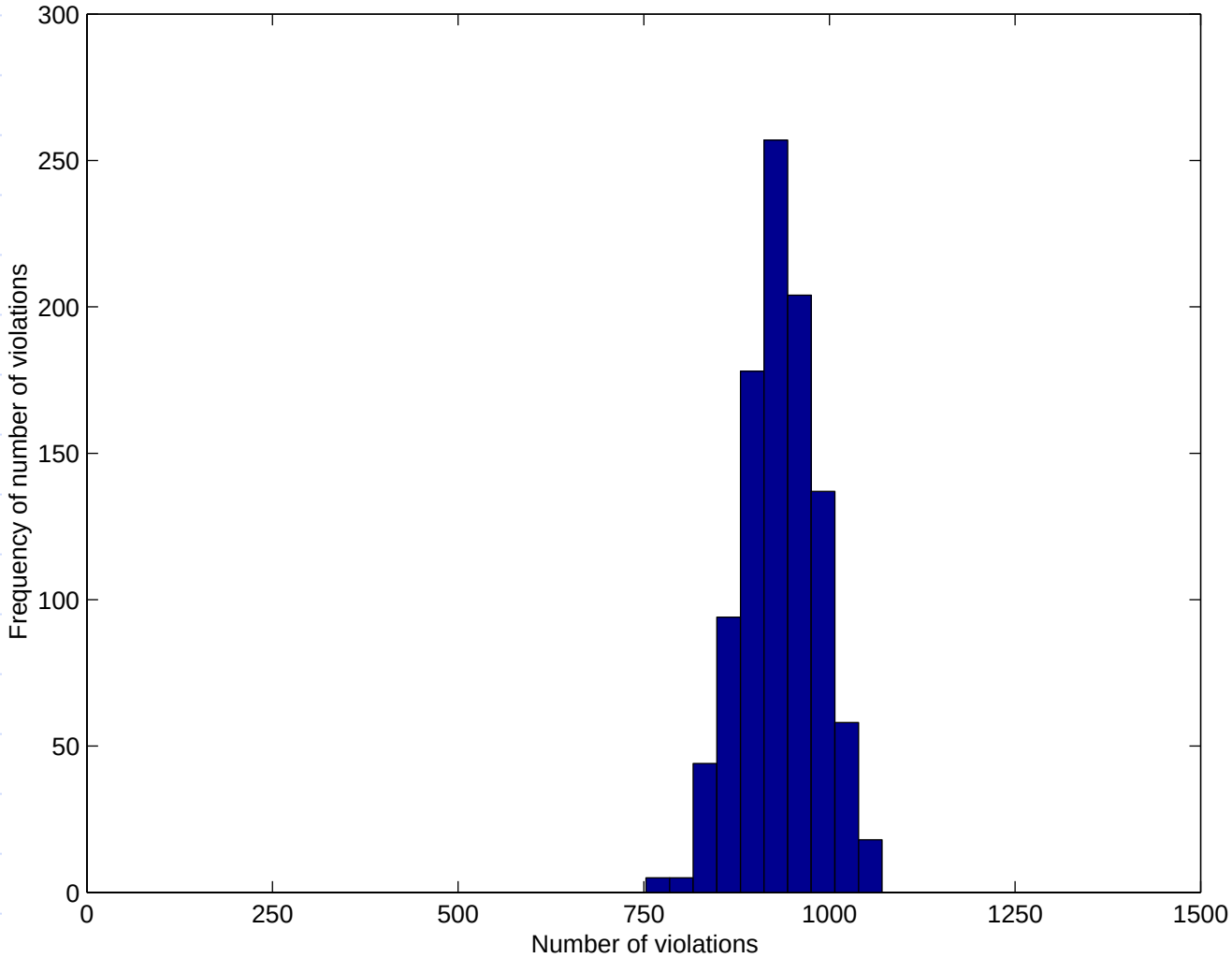
Results

Design	# Buffers	# Nodes	Violations	CPU Time
C1 (0.18μ)	616	4602	0	3.79
C2 (0.13μ)	588	3325	0	3.04

- Technique finds a *feasible* placement.
- CPU time is fast enough for iteration.
- Results were verified using detailed simulation.

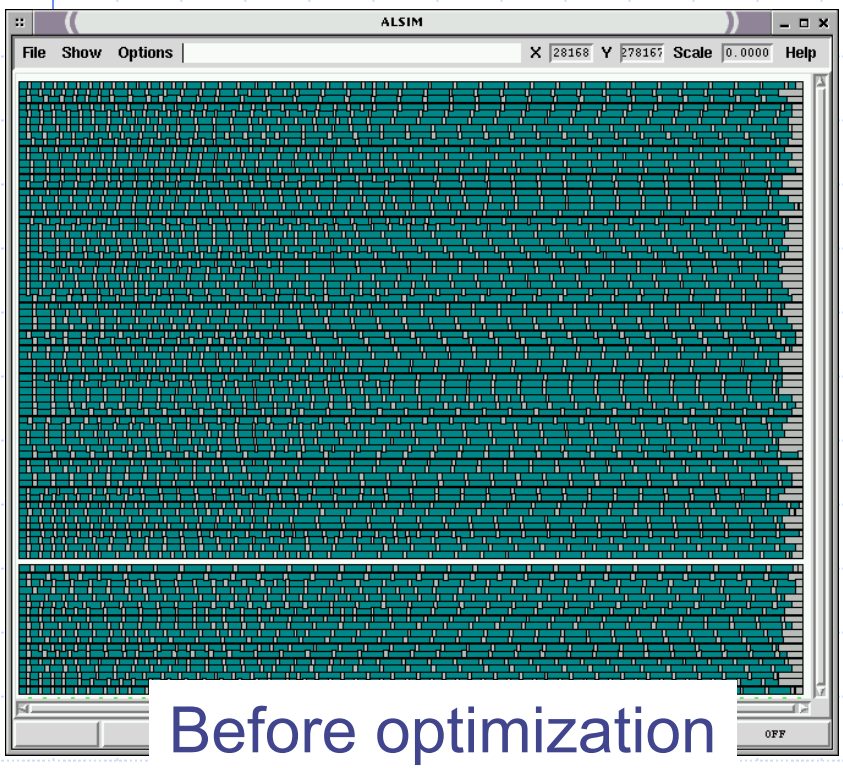
Is This an Easy Problem?

- Results of placing the I/Os randomly with 1.0% drop thresholds:

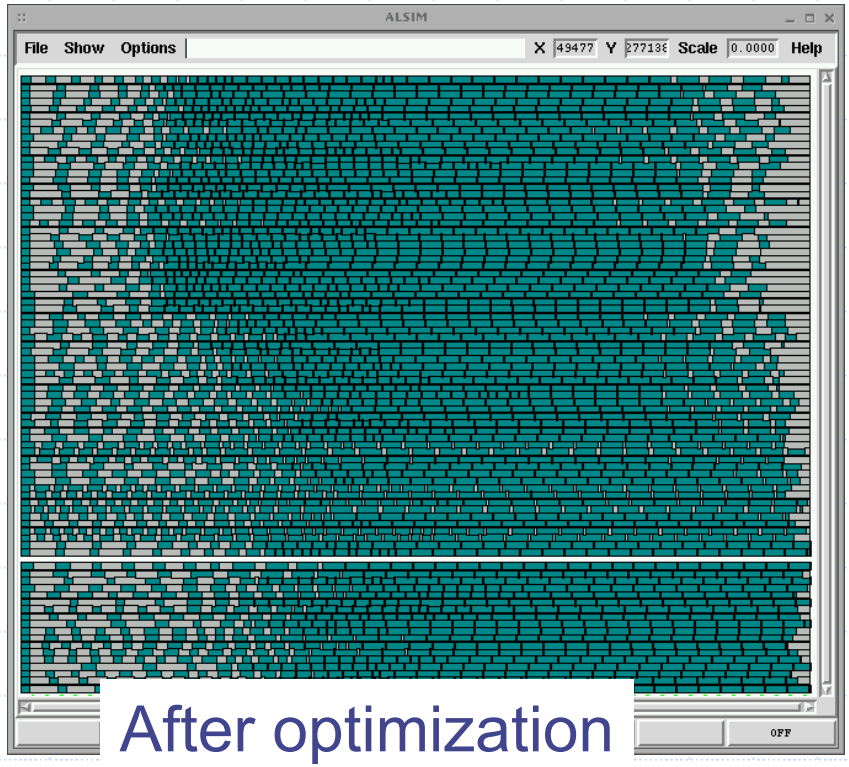


Decoupling Capacitance Sizing

- Upcoming DAC'02 paper (H. Su *et. al.*)



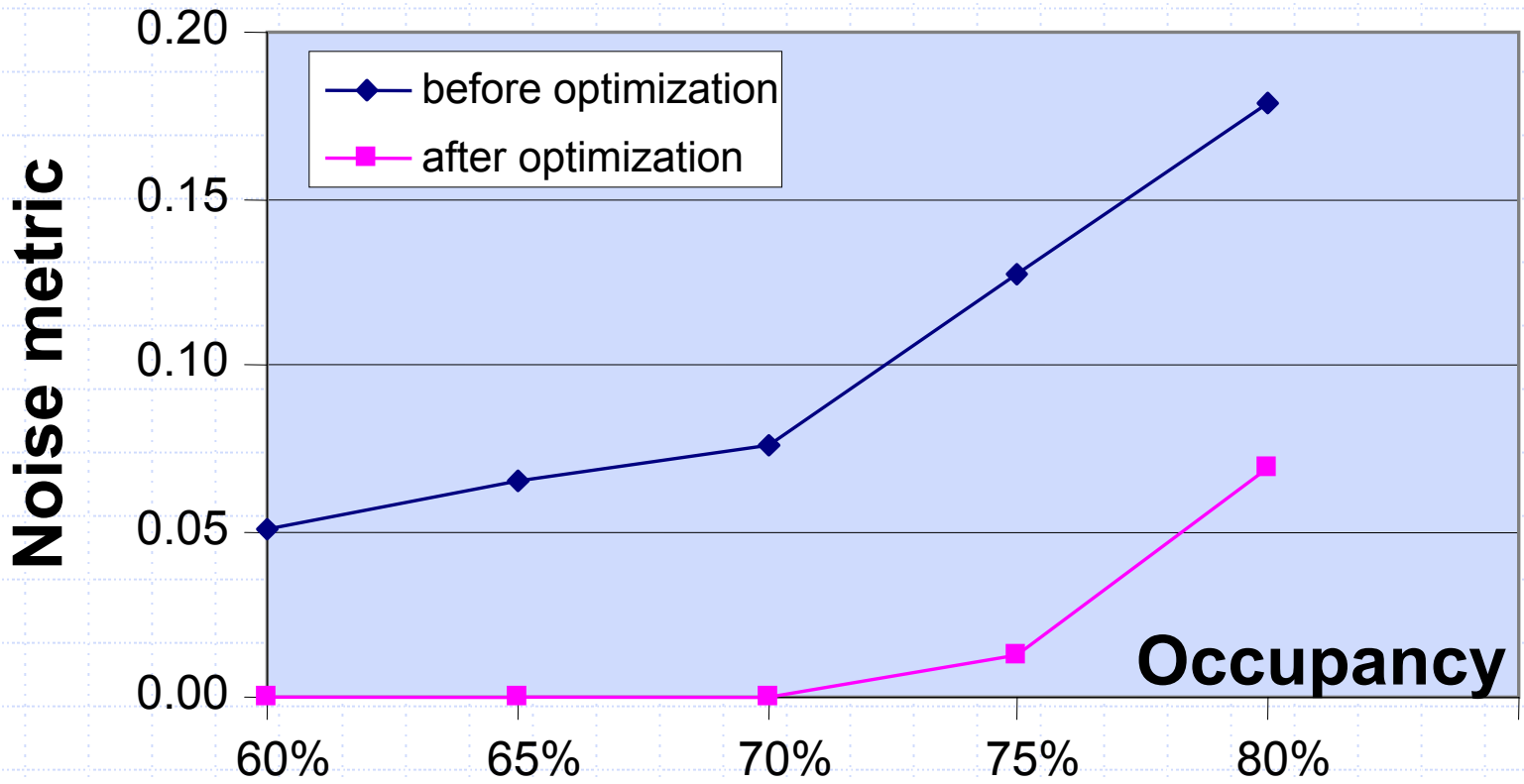
Before optimization



After optimization

Impact of Decap Sizing

- Decap optimization allows ~10% more circuit density!



Outline

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

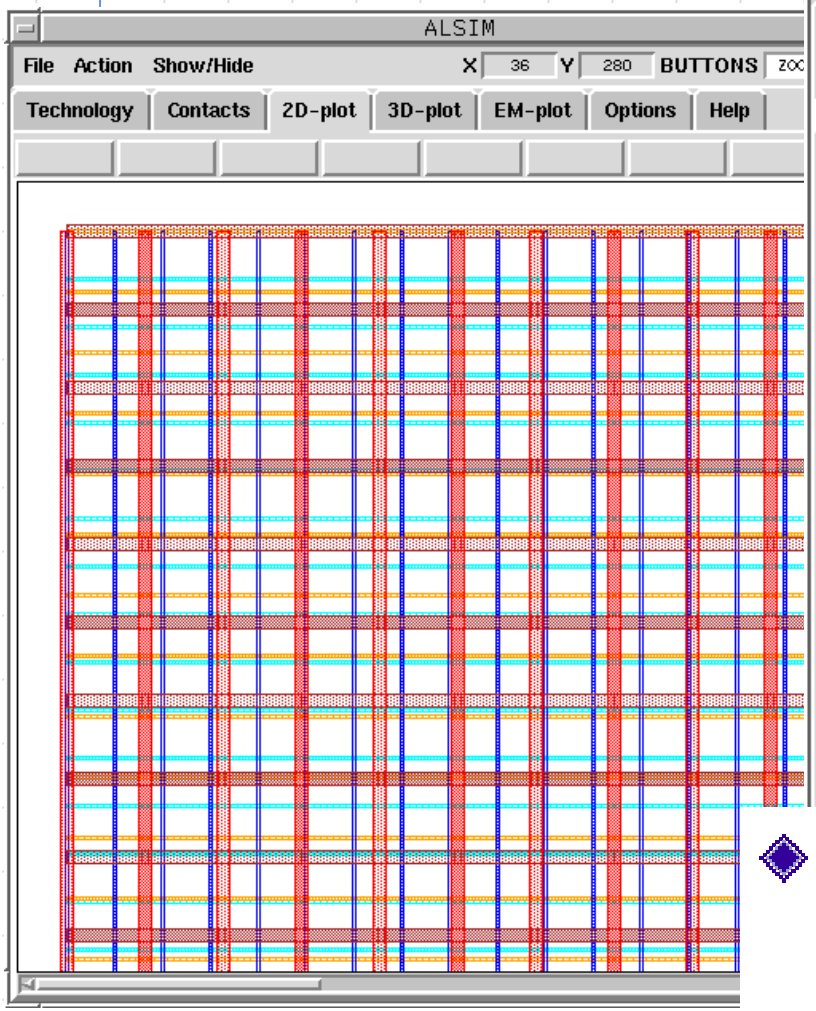
- Design realities and trends…
- Power Grid Planning.
- Power Grid Planning Examples.

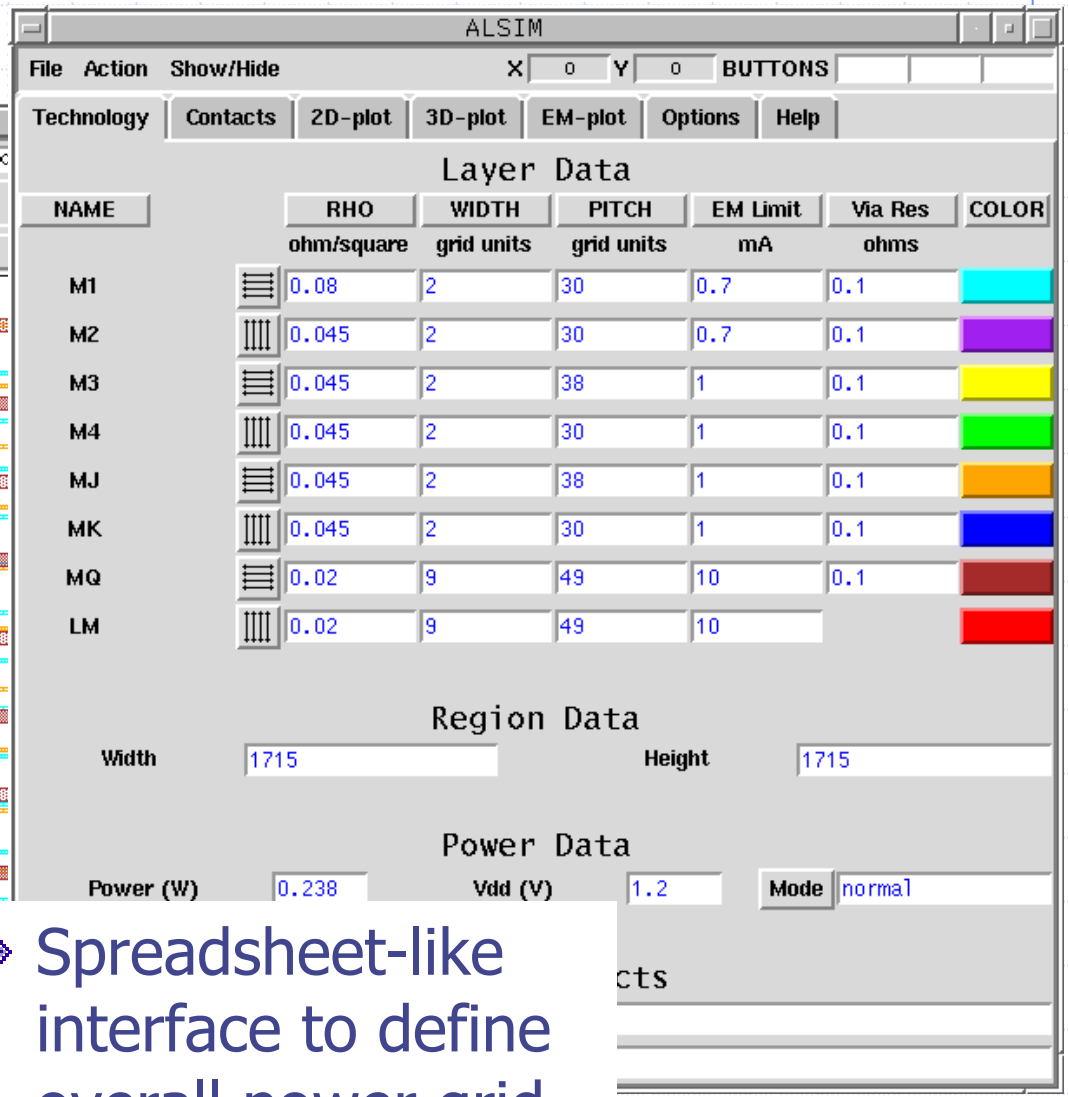
- Open issues and low-hanging fruit.

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

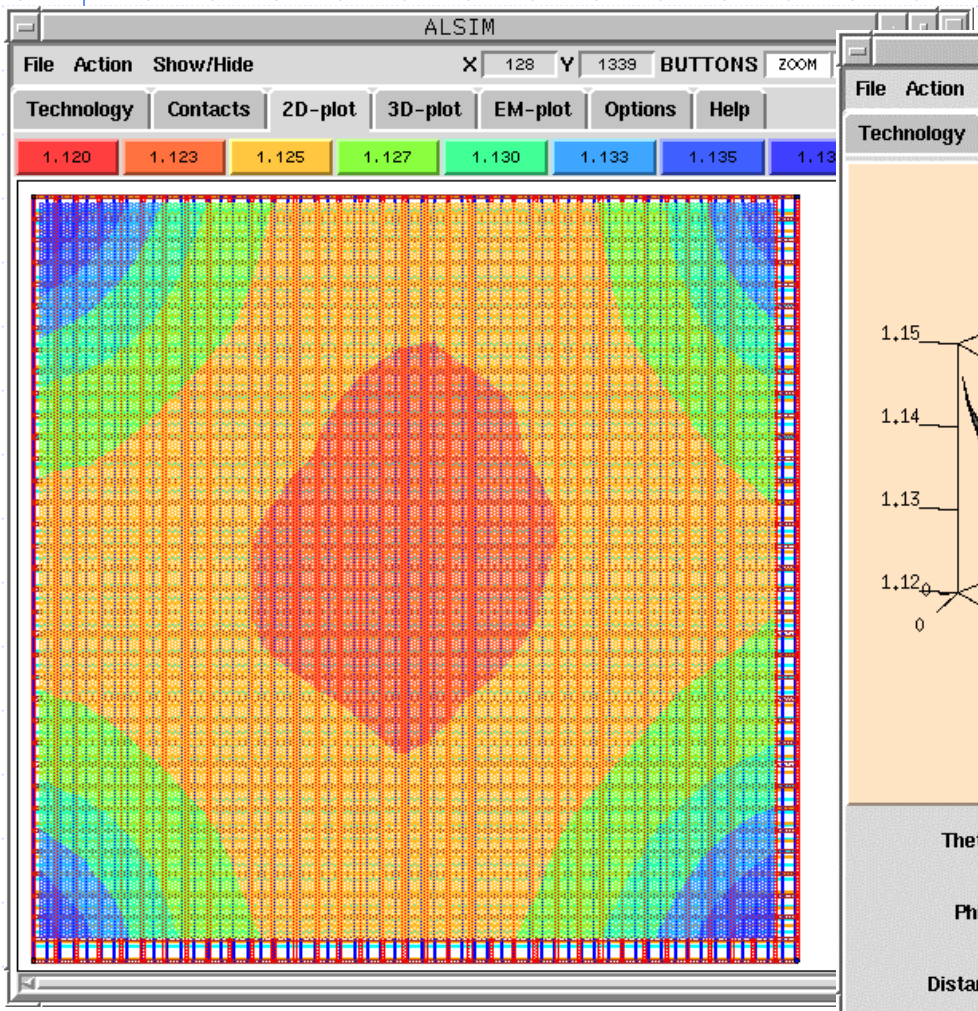


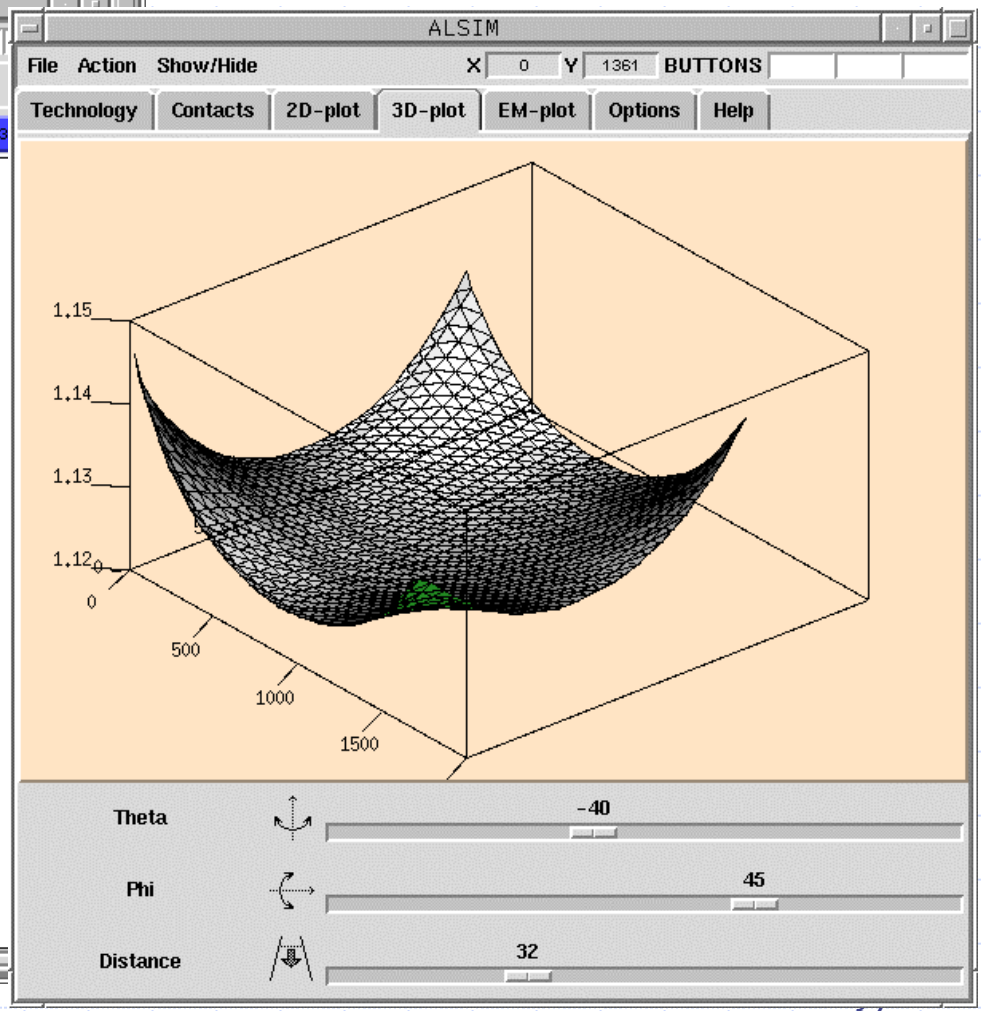


- Spreadsheet-like interface to define overall power grid.

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.

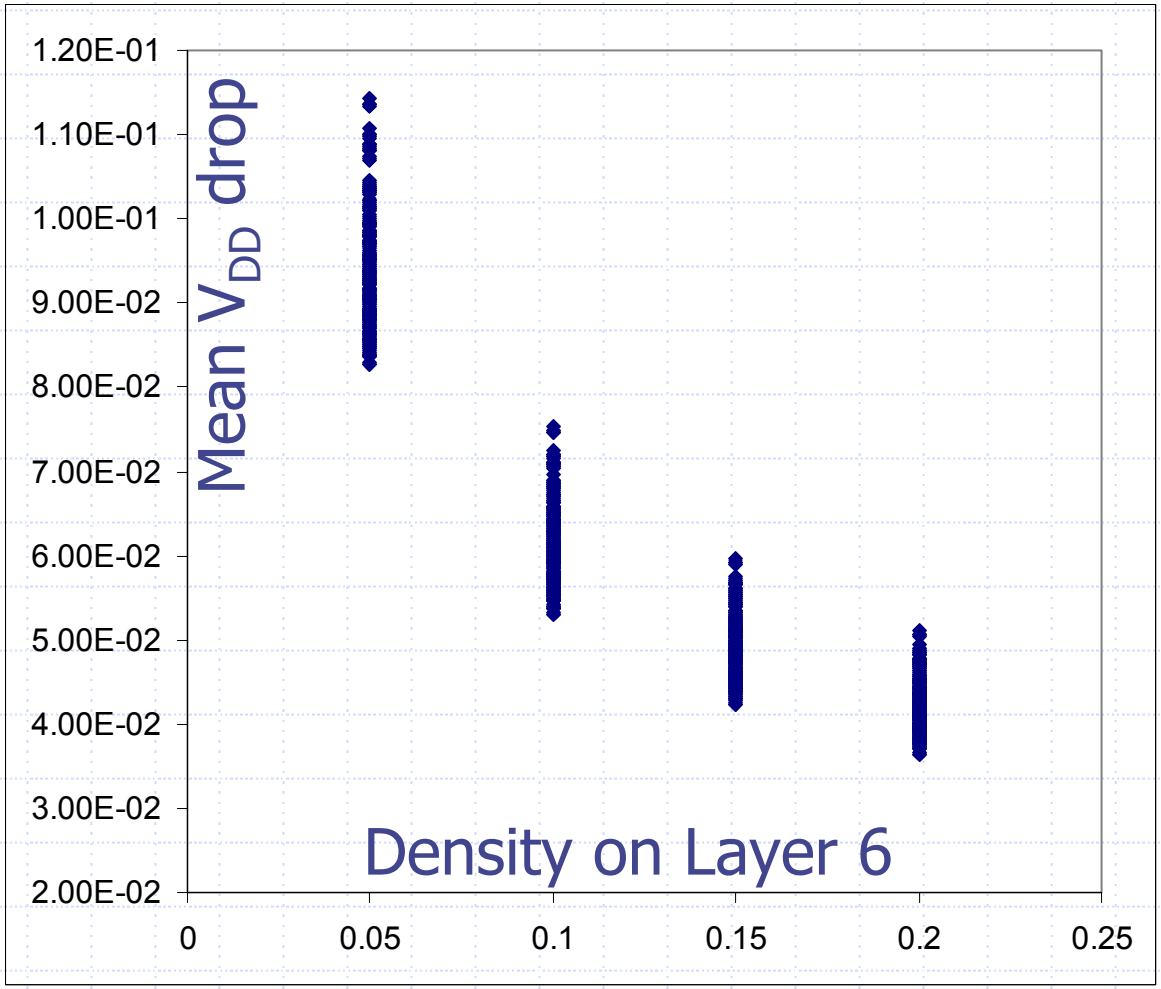
Experiment

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

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

- Measure VDD and GND net statistics.

Example of Results



Analysis of Results

- Linear regression of noise vs. layer densities.

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

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.