Thermal-Aware Clock Tree Design

Enrico Macii
EDA GROUP
POLITECNICO DI TORINO
Enrico.macii@polito.it

Outline

- Why temperature effects are important in DSM design

- Factors contributing to temperature increase

- Current densities – definitions and their implications

- Temperature-aware design methodologies

- Temperature effects in clock distribution networks

- Thermal-aware clock-tree generation

- Conclusions

The research team

- Massimo Poncino, Associate Professor

- Alberto Macii, Assistant Professor

- Ashutosh Chakraborty, PhD Student

- Prassanna Sithambaram, PhD Student

- Karthik Duraisami, Research Assistant

Why temperature effects are important in DSM

- Ever increasing chip power-density coupled with various other factors have contributed to large on-chip temperature gradients.

- Reliability is becoming a huge problem in current and future nanometric designs.

- Leakage power, which constitutes a major portion of the power consumption in nanometric designs, is exponentially dependent on temperature.

- Signal integrity due to temperature gradients in high performance ICs is becoming a major design problem to tackle.

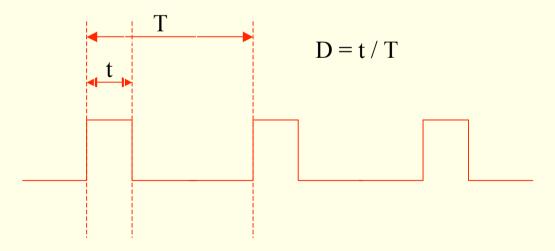# Factors contributing to an increase in temperature

- Aggressive interconnect scaling has resulted in higher current densities.

- Increase in the number of metal layers.

- Low-K dielectrics introduced in current silicon processes have very low thermal conductivity.

- Voltage does not scale in the same proportion as the rest of the geometries – higher power density.

- Different dynamic voltage scaling and clock-gating techniques have contributed to large temperature gradients on chip.

Current densities - Definitions and their implications

- Peak current density

- Average current density

- RMS current density

Current densities - definitions and their implications

■ Peak current density, Jpeak = Ipeak/A

■ Average current density = D * Jpeak

■ RMS current density = $\sqrt{D}$ * Jpeak

$D = t / T$

Current densities - definitions and their implications

Electromigration (EM)

- EM is the transport of mass in metals under an applied current density and is widely regarded as a major wear-out or failure mechanism of VLSI interconnects.

- TtF (Time to Failure) is exponentially dependent on the inverse of temperature.

- The EM lifetime reliability is proportional to the inverse of average current density.

Current densities - definitions and their implications

Joule Heating

- The power dissipated in the interconnect is dissipated as heat and this, in-turn, increases the temperature of the interconnect. This phenomenon is called "Joule heating".

- Joule heating varies as the square of the RMS current.

- So RMS current density determines the amount of self heating in interconnects as given by the following equation:

$$\Delta T_{self-heating} = (T_m - T_{ref}) = \frac{1}{T}\int_0^T I^2 R R_\theta \, \mathrm{d}t = I_{rms}^2 R R_\theta$$

Thermal-aware design methodologies

- Even though thermal modeling has received much attention from the scientific community, very little has been done in circuit design techniques to reduce hazards caused by temperature gradients in the ICs.

- Thermal-aware design must be a part of the design flow for future 3D ICs.

- Thermal-aware placement, thermal floor-planning, thermal-aware clock distribution are some of the areas that need to be explored.

Temperature effects on clock distribution network
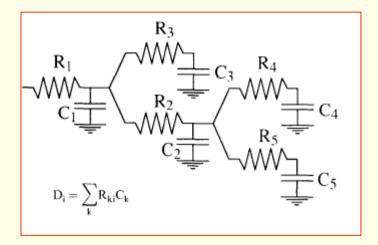
Thermal gradients – Sources

- Power reduction techniques such as dynamic power management, clock-gating, operand isolation, multi V_{dd} and V_{th} devices, sleep-transistor insertion, transistor sizing induce temperature gradients on the substrate.

- With ever decreasing feature sizes, the global metal layers on which the clock signal is routed are getting closer to the substrate.

- Temperature gradients in clock networks may be induced due to self heating or thermal coupling from the substrate or metal layers underneath the clock.

Temperature effects on clock distribution network

Effects of temperature gradients

- Clock skew induced by temperature gradients is no longer negligible.

- Buffer insertion in clock networks have to be revisited to account for temperature effects.

Temperature effects on clock distribution network

Clock Skew (Basics)

- Delay in a clock tree is modeled by the Elmore delay model.

- With temperature gradients, R is no longer a constant, but it becomes temperature dependent.

- $R(x) = R0(1+ \beta\ T(x)\ )$

- Skew is no longer zero and is now dependent on temperature profile of the clock trace.
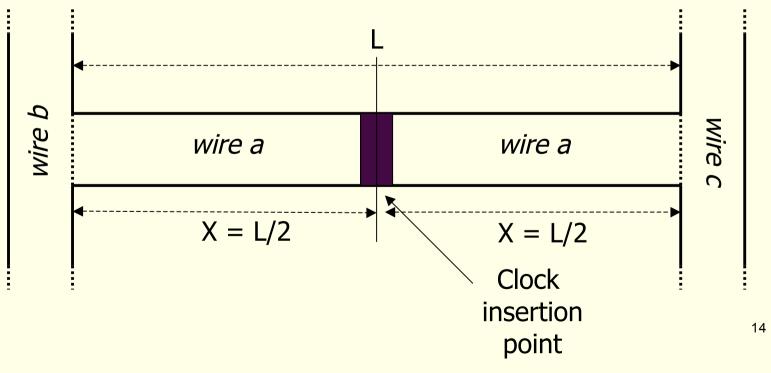
$$D_i = \sum_k R_{ki} C_k$$

- This can be easily seen in clock gating circuits where a portion of the circuit is deactivated for a certain period of time and then is activated as dictated by the clock gating algorithm.
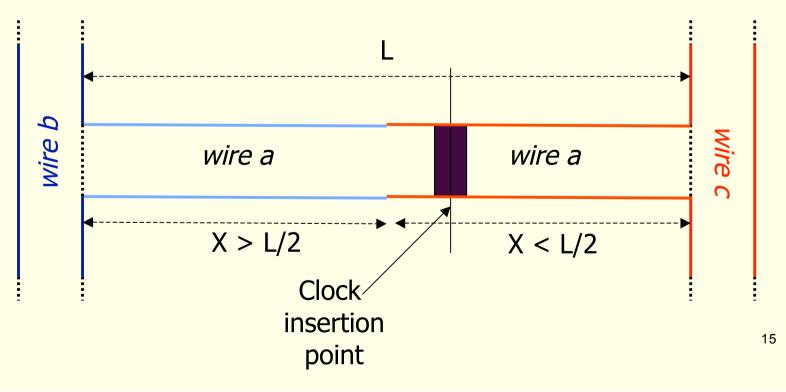
Temperature effects on clock distribution network

- ■ Assumption:
  - ■ Uniform thermal profile
- ■ Target:
  - ■ Zero-skew clock tree

L

wire b

wire a   wire a

wire c

X = L/2   X = L/2

Clock
insertion
point

14

Temperature effects on clock distribution network

- **Assumption:**
  - Thermal profile linearly increasing towards wire **c**
- **Target:**
  - Zero-skew clock tree

15

Thermal-aware clock-tree design

- Solutions proposed for zero-skew clock trees and simple thermal profiles (e.g., ICCAD-05 paper by D. Z. Pan, UT Austin).

Our Approach

- Capable of dealing with non-zero skew clock trees

- Significant reduction in worst case clock skew.

- Minimum wirelength penalty.

- Improves on a BST-DME based algorithm.

- Robust to process variations as it is minimally intrusive and maintains the *goodness* of the original tree.
  - Starting point is an existing clock tree.

Thermal-aware clock-tree design

Solution Region Based Optimization (SRBO)

- Works on Bounded Skew Trees as against Zero Skew Trees.

- Two stage algorithm that builds Solution Regions for all nodes during a *bottom-up* phase and then embeds the node during the subsequent *top-down* phase.

Algorithm 1 Complete Flow

1: **INPUT:** Original BST/DME generated Tree TR with Bound B
2: **INPUT:** Temperature distribution across the chip, $T(X)$
3: $TR_T \leftarrow$ Modified Tree
4: $B_T \leftarrow$ Modified Skew
5: Build_Bottom_Up$(TR,T(X))$
6: $TR_T =$ Embed_Top_Down(TR)
7: Update_Tree_Characteristics(TR_T)
8: $B_T =$ Recalculate_Skew$(TR_T,T(X))$
9: **OUTPUT:** Temperature aware modified tree TR_T with skew B_T and wirelength WL_T

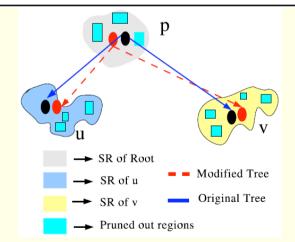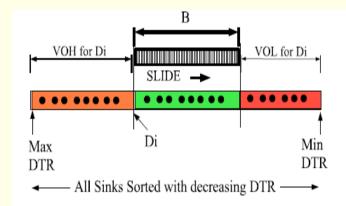- SR of Root
- SR of u
- SR of v
- Pruned out regions
- Modified Tree
- Original Tree

Illustration showing the modified tree after temperature aware re-embedding with Solution regions

Thermal-aware clock-tree design

Critical Path Based Optimization (CPBO)

■ Fast heuristic that performs temperature aware modifications only on *critical* paths.

■ Less computation time as compared with the *Solution Region Based Optimization* algorithm.

■ The algorithm sorts the *DelayTillRoot (DTR) values* for all sinks in the design.

■ Sinks that fall outside of the acceptable value of skew (B) are isolated and moved around in the Manhattan space to yield a less worst case skew for the non-uniform thermal profile.
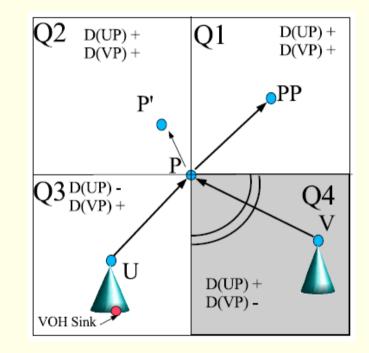
Thermal-aware clock-tree design
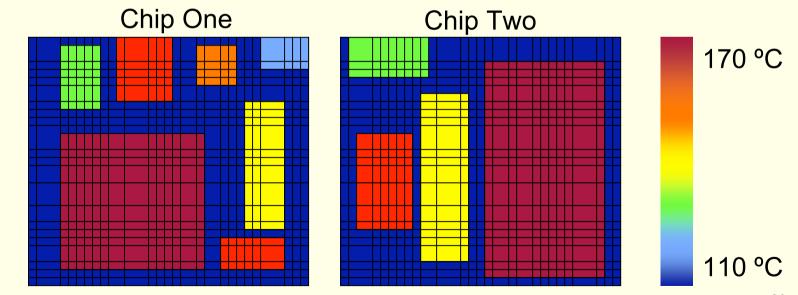
Critical Path Based Optimization
(CPBO)

Quadrant Optimization

- To reduce computation time in arriving at the best possible embedding location for a particular Steiner point.

- Given U is a critical sink (delay from P to U being greater than delay from P to V), the final embedding is narrowed down to Quadrant 4.

- Reduces computation time by one-fourth

Thermal-aware clock-tree design: Results

- Standard benchmarks (ICCAD-91 paper by R. Tsay).
- Thermal profiles:
  - Linearly increasing along chip width.
  - Chip One.
  - Chip Two.

Chip One   Chip Two

170 ºC

110 ºC

Thermal-aware clock-tree design: Results

Results for SRBO

Avg. skew
savings: 56%

Avg. wiring penalty:
< 1%

Bench Mark	Therm Skew(ps)	SRBO Algorithm		
		Optimized(ps)	Savings Linear	Wire-Penalty(%)
(A)	(B)	(C)	(D)	(E)
P1	121	107	66.66	0.41
P2	171	138	46.47	0.36
R1	181	149	39.50	0.69
R2	365	184	68.30	0.84
R3	271	149	71.34	0.93
R4	663	318	61.27	1.04
R5	1268	723	46.66	1.09
		Profile One		
P1	168	121	69.11	0.79
P2	386	203	63.98	0.87
R1	246	165	55.47	0.71
R2	324	181	63.83	0.93
R3	891	463	54.10	0.91
R4	1494	794	50.21	0.97
R5	2785	1578	44.95	1.04
		Profile Two		
P1	128	109	67.85	1.02
P2	325	237	39.11	0.75
R1	164	125	60.93	1.01
R2	240	159	57.86	1.14
R3	397	221	59.25	1.07
R4	1532	754	54.32	1.01
R5	2314	1201	50.27	0.98

Thermal-aware clock-tree design: Results

Results for CPBO

Avg. skew
savings: 51%

Avg. wiring penalty:
< 1%

CPBO algorithm is
4 times faster than
SRBO

		CPBO	Algorithm	
Bench Mark	Therm Skew(ps)	Optimized(ps)	Savings Profile	Wire-Penalty(%)
(A)	(B)	(F)	(G)	(H)
P1	121	109	57.14	0.27
P2	171	141	42.25	0.21
R1	181	154	33.33	0.16
R2	365	194	64.52	0.63
R3	271	154	68.42	0.62
R4	663	340	57.37	0.52
R5	1268	784	41.43	0.25
		One		
P1	168	127	60.29	0.36
P2	386	214	60.13	0.42
R1	246	174	49.31	0.33
R2	324	189	60.26	0.43
R3	891	487	51.07	0.35
R4	1494	823	48.13	0.27
R5	2785	1622	43.31	0.29
		Two		
P1	128	112	57.14	0.34
P2	325	247	34.66	0.23
R1	164	130	53.12	0.41
R2	240	168	51.42	0.38
R3	397	227	57.24	0.32
R4	1532	762	53.77	0.38
R5	2314	1284	46.52	0.27

Conclusions

- Temperature effects can no longer be considered trivial and hence design techniques and accurate modeling are absolutely necessary to build robust chips.

- Clock trees are even more susceptible to temperature gradients across the chip as they span the entire die, making thermal aware clock-tree construction an absolute necessity.

- We have proposed two algorithms for thermal-aware clock tree re-design. Results, although they have come out of the oven thew days ago (DATE 2006 submission), are very promising.