Power-Aware and Temperature-Aware Circuit Design

Mircea R. Stan

ECE Department, University of Virginia
EECS Dept., Univ. of California at Berkeley

Charlottesville, VA 22904
Email: mircea@virginia.edu

UVa program in VLSI

High-Performance Low-Power VLSI Lab

- Currently:
  - 3 PhD students
- Alumni:
  - 1 (PhD) at IBM
  - 1 (PhD) at Bay Area startup
  - 1 (PhD) at Intel
  - 3 (MS) at Intel
  - 2 (MS) at MIPS
  - 1 (MS) at NY startup
- Interests: high-performance low-power circuits, mixed-mode circuits, embedded systems, nanoelectronics
- Visiting faculty: UC Berkeley (2004-2005), Intel (2002, 1999), IBM (2000)

Berkeley Wireless Research Center (BWRC)

- PicoRadio project, very low voltage circuit design

Overview

- Motivation
- Figures of Merit (FOM)
- Temperature-Adaptive Circuits (TAC)
- Temperature-Aware Computing (HotSpot)
- Conclusions and future work

Levels: circuit, logic, architecture

Work in progress!

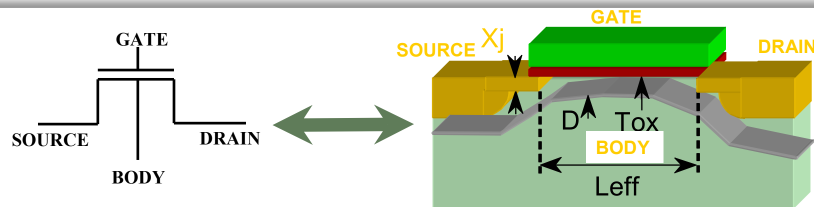
Overview

- Motivation
- Figures of Merit (FOM)
- Temperature-Adaptive Circuits (TAC)
- Temperature-Aware Computing (HotSpot)
- Conclusions and future work

Motivation

- Power: first-order design constraint
  - max power consumption: limits power delivery
  - sustained power dissipation: limits thermal design/packaging
  - average active power and idle power consumption: limit battery life, etc.
- Power-aware design:
  - maximize performance for given power
- Low-power design:
  - minimize power for required performance
- Temperature-aware design:
  - performance, power, reliability: function of T
  - T function of power density, ambient T
  - maximize performance for given thermal envelope

Technology Scaling



Dimensions scale down by 30%	Doubles transistor density
Oxide thickness scales down	Faster transistor, higher performance
Vdd & Vt scaling	Lower active power

Technology has scaled well, will it in the future?

Source: Shekhar Borkar, Intel Research

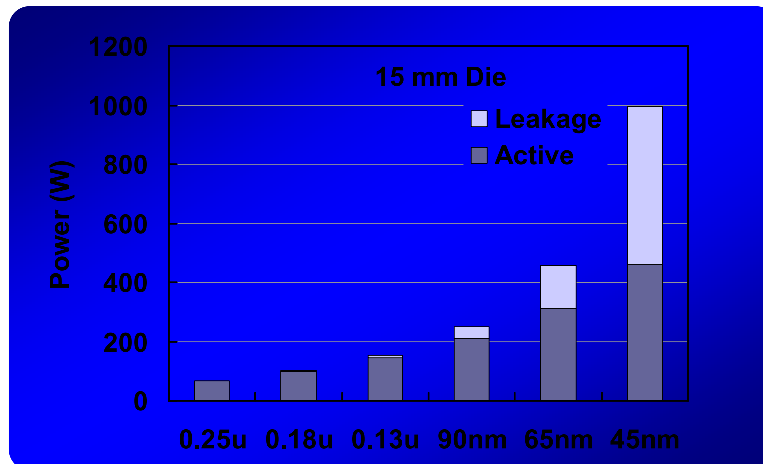
ITRS Projections

Year	2003	2006	2010	2013	2016
Tech node (nm)	100	70	45	32	22
Vdd (high perf) (V)	1.0	0.9	0.6	0.5	0.4
Vdd (low power) (V)	1.1	1.0	0.8	0.7	0.6
Frequency (high perf) (GHz)	3.1	5.6	11.5	19.3	28.8
	Max power (W)				
High-perf w/ heatsink	160	180	218	251	288
Cost-performance	85	98	120	138	158
Hand-held	3.2	3.5	3.0	3.0	3.0

ITRS 2001

- **These are targets**
- **Power-density problem is getting even worse**
- **Intel papers suggest that in the 45-75W range, cooling costs $1/W; but then rate of increase goes up: $2, $3/W, probably more!**
  (Borkar, IEEE Micro '99, Gunther et al, ITJ '01)

The Power Crisis

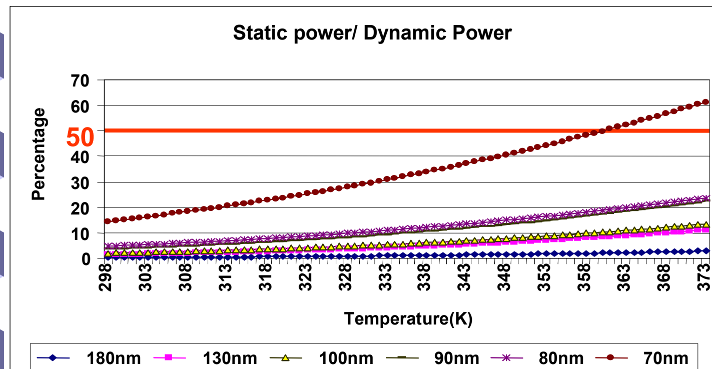


Business as usual is not an option

Source: Shekhar Borkar, Intel Research

Leakage Power

- The fraction of leakage power is increasing exponentially with each generation
- Also exponentially dependent on temperature



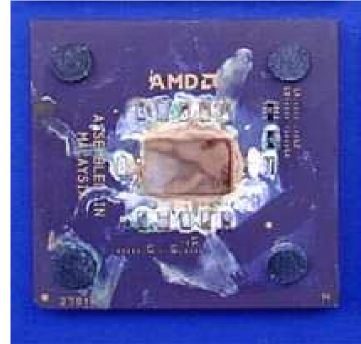
Source: Sankaranarayanan et al, University of Virginia

Future of Scaling?

- Constant Voltage scaling: old
  - Power and tox problem
- Constant Field scaling (scale Vdd): now, but still Constant Temperature
  - Leakage problem
- Scaling as we know it has stopped!
- Likely: future also scale Temperature
  - Only applicable for servers
  - Constant Ion/Ioff
  - In theory solves leakage problem
  - Need efficient active cooling solutions
  - Not for portables obviously!

Power and temperature are BAD

- and can be **EVIL**



Source: Tom's Hardware Guide
http://www6.tomshardware.com/cpu/01q3/010917/heatvideo-01.html

Overview

- Motivation
- Figures of Merit (FOM)
- Temperature-Adaptive Circuits (TAC)
- Temperature-Aware Computing (HotSpot)
- Conclusions and future work

Power-aware figures of merit (FOM)

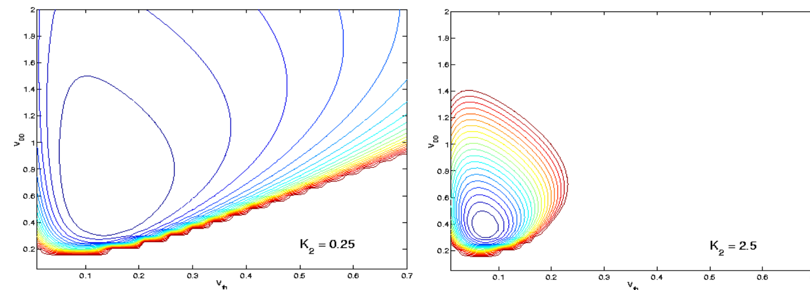
- **Power (P):** battery time (mobile)
  packaging (high-performance)
- **Energy (PD):** battery life (mobile)
  fundamental limits (kT)
- **Energy-delay (PD^2):**
  performance and low power
- **Energy-delay^2 (PD^3):**
  emphasis on performance

Designers: first optimize architecture (PD^3), then optimize voltages (PD^2)

2-D like "old" VLSI complexity (A, AD, AD^2)

Refs: A. Martin et al. "Design of an Asynchronous MIPS R3000", ARVLSI'97
Gonzales etal. "Supply and threshold voltage scaling for low power CMOS", JSSC, Aug.1997
J. Ullman, "Computational aspects of VLSI", CS Press, 1984

Optimum voltages for PT^2



Optimum Vdd and Vth:

$$K_2 = \frac{a}{g L_d V_T^2} \qquad e^{-n} = 1.5 K_2 \cdot n^{0.5}$$

$$V_{dd} = \frac{3}{3-\alpha} V_{th} + \frac{3\alpha}{3-\alpha} \eta V_T$$

$$\frac{n + 2\alpha - 3}{e^n \cdot (\alpha \cdot (n+3) \cdot \eta V_T)^{\alpha} \cdot (3-\alpha)^{1-\alpha}} = K_2$$

$$n = \frac{V_{th}}{\eta V_T}$$

Optimum ratio of AC to DC power (for PD^2):

$$\left(\frac{P_{AC}}{P_{DC}}\right)_{opt} = \frac{PT_{AC}^2}{PT_{DC}^2} \approx \frac{K_2 \cdot n^{1.5}}{e^{-n}} = \frac{2}{3} \frac{V_{th}}{\eta V_T}$$

Optimum Vdd, Vth – independent of technology

- For given circuit: same values for all technologies
- Still: function of activity, logic depth, temperature

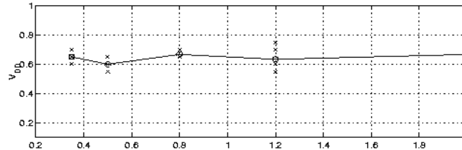
Quantitative: Vth not scaling!

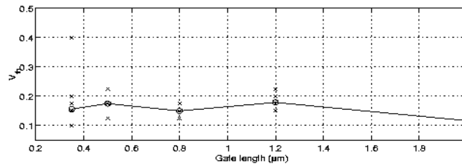
$$V_{DD} = 3V_{th} + 6nV_T$$

$$\frac{a}{gL_d} = \frac{V_{th} + nV_T}{(V_{th} + 3nV_T)^2} \frac{nV_T}{4} e^{-\frac{V_{th}}{nV_T}}$$

SIMULATIONS RESULTS

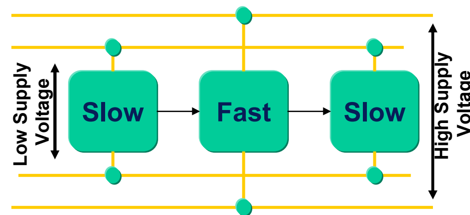
L μm	Manuf.	Model	PT^2 $10^{-20} J$	V_{DD} (V)	V_{th} (V)
2	Orbit	Level3	4.38	0.6	0.15
		Level2	5.31	0.65	0.125
		Bsim3	5.5	0.75	0.075
1.2	AMI	Bsim3	2.19	0.75	0.2
			1.31	0.55	0.225
			1.73	0.7	0.175
		Level3	1.08	0.6	0.15
			0.94	0.55	0.15
			1.15	0.65	0.175
0.8	HP	Bsim3	0.60	0.65	0.175
			0.65	0.7	0.15
			0.66	0.65	0.125
0.5	HP	Bsim3	0.37	0.6	0.225
			0.4	0.65	0.175
			0.20	0.55	0.125
0.35	TSMC	Bsim3	1.04	0.65	0.1
			1.05	0.65	0.175
	HP		0.99	0.7	0.2
			0.37	0.65	0.4
			1.02	0.6	0.15





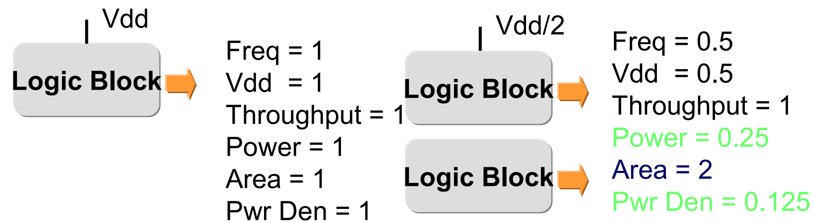
Ref: A. Forestier et al. "Limits to voltage scaling...", SBCCI, Sep. 2000

Multi-Vdd



Multiple Supply Voltages

Replicated Designs

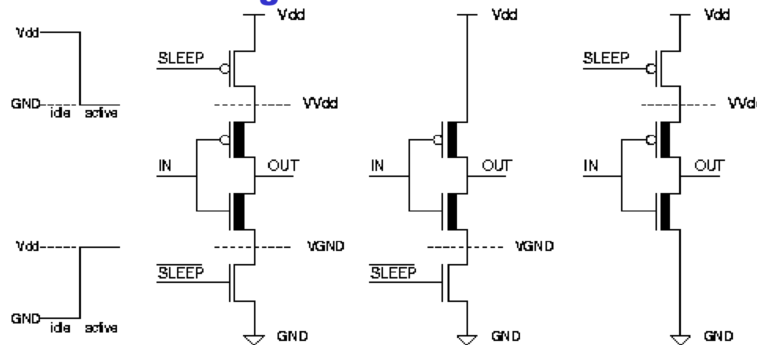


Source: Shekhar Borkar, Intel Research

Multiple Vth - MTCMOS

High activity: low Vth, Low activity: high Vth

Both low and high: multithreshold - MTCMOS



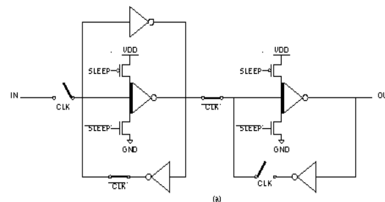
Sleep mode: power supply is gated

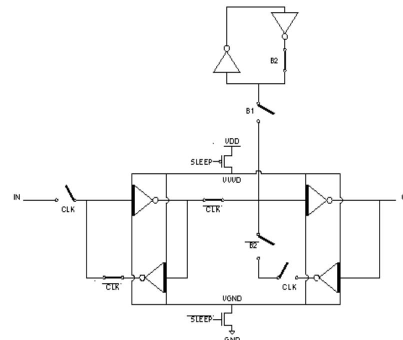
Many other possible variations!

Challenge: keeping the **state** during sleep

Previous MTCMOS flip-flops

- Duplicate structure
- Separate structure: "balloon" circuit





- Intermediate performance
- Area overhead
- Distributed gating

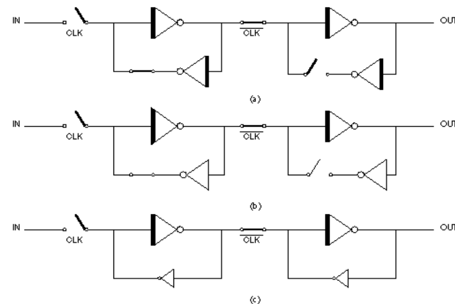
- Area overhead
- Complex control
- Centralized gating (virtual supply rails)

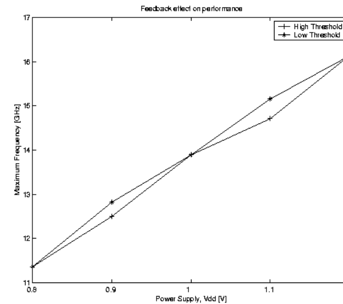
Ref: Mutoh et al. "MTCMOS…", IJSSC

Novel MTCMOS flip-flops

New ideas:

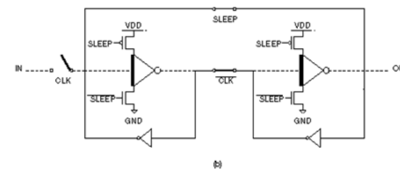
- **Small area overhead: reuse (high-Vth) transistors**
- **Small performance penalty: simple circuit**
- **Use outer feedback**



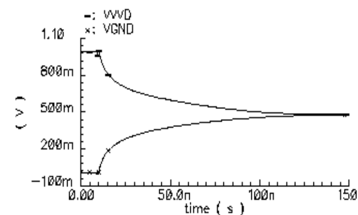


MTCMOS with outer feedback (MTOF)

- **Active mode: same**
- **Sleep mode: outer feedback and reuse**
- **Low area overhead**
- **High performance**
- **Simple control**



Still: distributed gating to avoid sneak paths

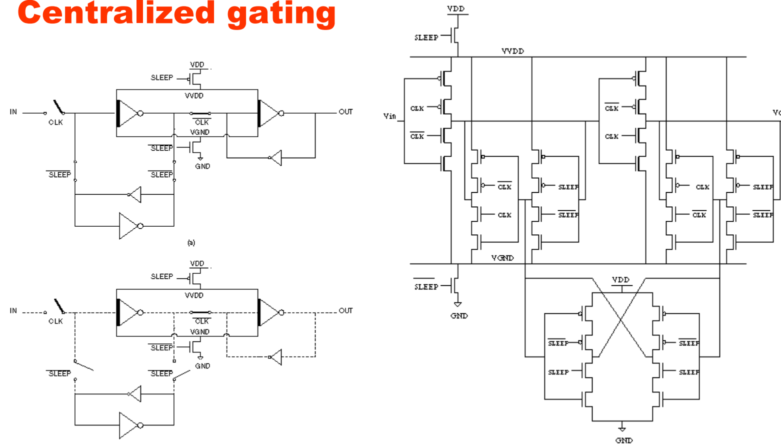


Ref: M. Stan et al. "MTOF", US patent

MTOF static flip-flops

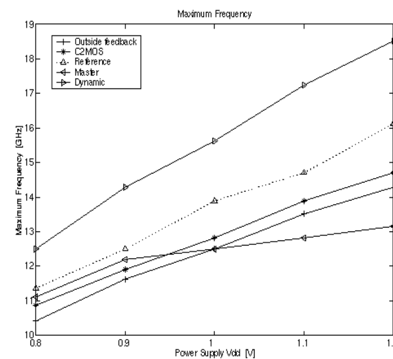
Different topologies:

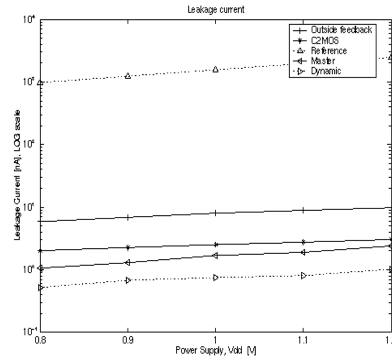
- Master-only
- C2MOS: active feedback
- Centralized gating



Performance

Flip-flop Type	f_{max} [GHz]	CLK-OUT Delay [ps]	Set-up Time [ps]	Leakage Power [nW]	Area: Total Width [um]	Type	Complexity
Reference	13.90	35.26	36.74	1570	15	no gating	reference
Duplicated Structure	11.11	47.00	43.00	6.90	15.75	distributed	very low
'balloon'	13.15	41.90	34.10	4.30	20.25	cenralised	very high
Outside Feedback	12.73	45.8	32.75	8.02	14.25	distributed	very low
C^2MOS	12.82	54.00	24.00	2.52	21	centralised	very low
Master	12.67	39.20	39.92	1.69	15.75	centralised	very low
Outside Feedback MVCMOS	12.20	47.90	34.07	8.8e-3	14.25	distributed	very low
C^2MOS MVCMOS	12.50	47.20	33.15	8.4e-3	21	centralised	very low
Dynamic	15.62	35.60	28.40	0.75	15	centralised	low





Summary power-aware FOM

- Gigascale transistor integration capacity will continue to be available—Power, Power Density and Energy are the barriers
- Improve design efficiency for desired FOM
- Multi—everywhere: multi-Vdd, multi-Vth, dynamic voltage scaling (DVS), automatic body-bias (ABB), multi-threading, etc.
- Exploit integration capacity to deliver performance in power/cost envelope

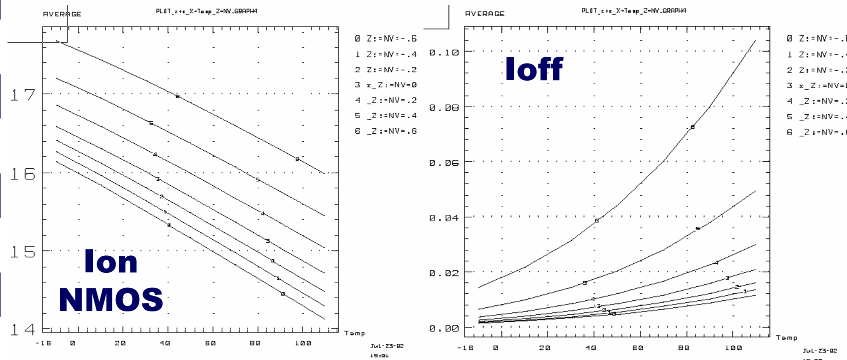
Overview

- Motivation
- Figures of Merit (FOM)
- **Temperature-Adaptive Circuits (TAC)**
- Temperature-Aware Computing (HotSpot)
- Conclusions and future work

Thermal effects

Temperature (Berkeley PTM 70nm CMOS):

- **Transistor threshold and mobility**
- **Subthreshold leakage, gate leakage**
- **Ion, Ioff, delay**

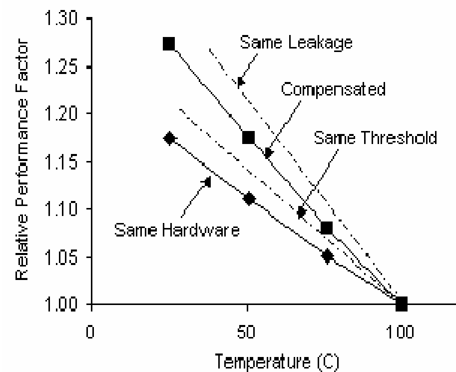


High performance circuits

- **Robustness constraint: sets Ion/Ioff ratio**
- **Robustness and reliability: Ion/Igate ratio**
- **70nm CMOS, 1.2V, 110ºC**
- **Ion/Ioff ~ 1000**
- **Ion/Igate ~ 10000**

Idea: keep ratio constant with T

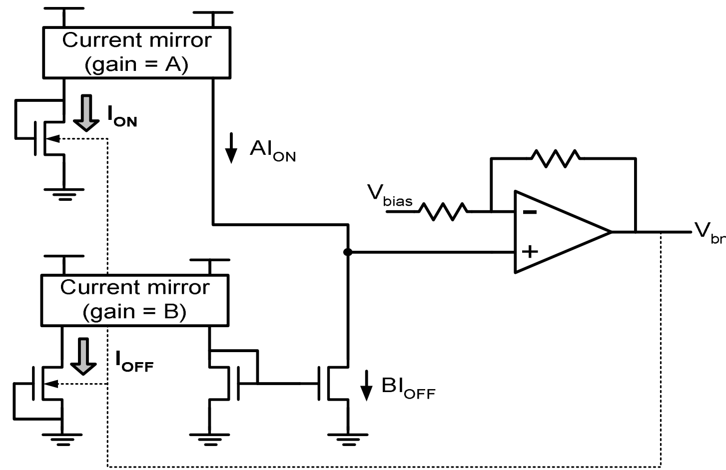
Trade leakage for performance



Ref: "Ghoshal et al. "Refrigeration Technologies...", ISSCC 2000
Garrett et al. "T3...", ISCAS 2001

Adaptive Ion/Ix control

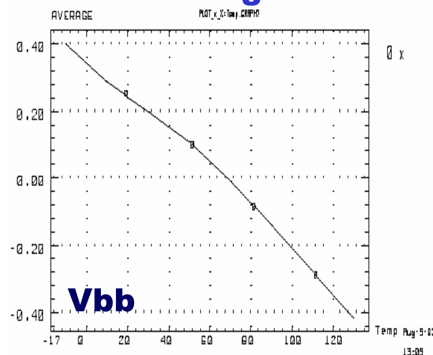
Ion/Ioff = B/A = ct. through ABB

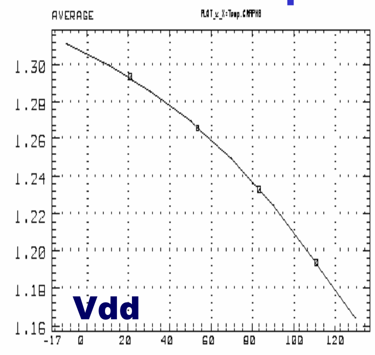


Temperature-aware circuits (TAC) patent pending (2002)

Resulting voltages

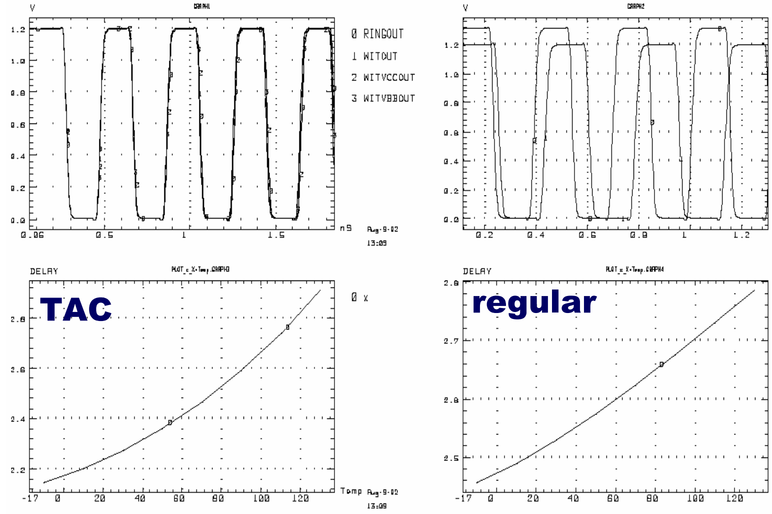
- Wide range: -.4V < Vbb < .4V; 1.2V < Vdd < 1.3V
- Almost linear
- Robust to inter-die parameter variations
- Needs trimming for setpoint
- Margin for intra-die parameter variations
- Active cooling or natural thermal landscape





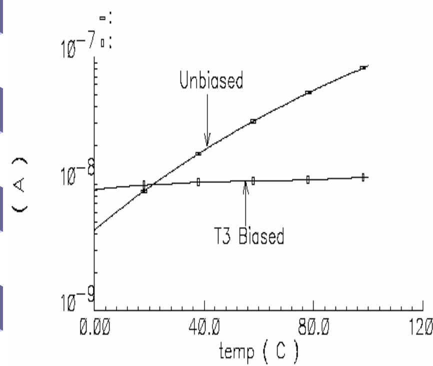
Resulting performance

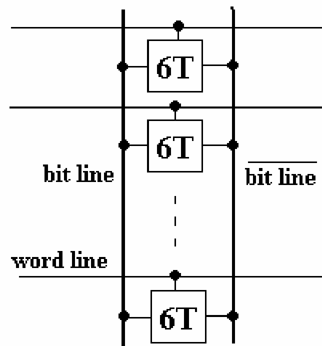
- 25% extra performance (110°C to 0°C) – only NMOS
- 13% from low temperature alone



SRAM Read time

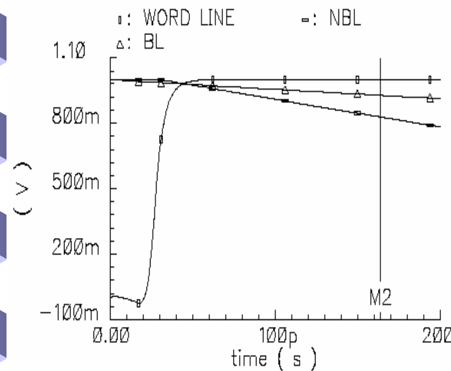
- Same circuit, different application
- 6T SRAM memory: "reverse application" (heating)
- 70nm process (200mV threshold)
- Zero biasing at low temperature

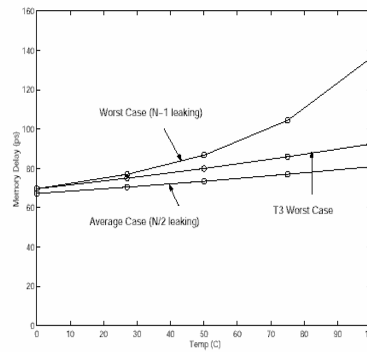




SRAM bit-line sensing

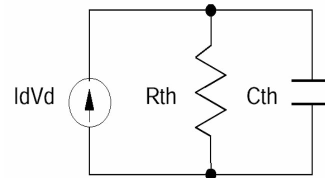
- Differential sensing (100mV bitline difference)
- 128 cells per bit line
- Faster read even if higher RBB, smaller Ion

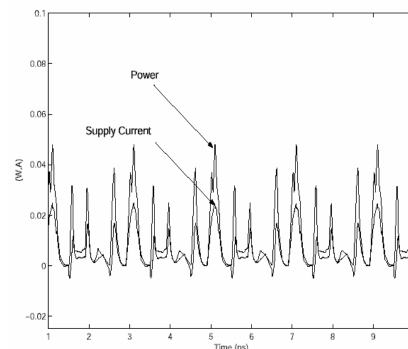




Electro-thermal circuit cosimulation

- Temperature modeling
  - MOS parameters
  - time scale (ms)
- Self-heating transistor models (BSIMPD)
- Thermal equivalent circuits
- Need convergence of BSIM3v3 (industry standard, BSIM4 (Igate), BSIMPD (temperature pin)
- More work needed!





Summary TAC

Project started at UVa, funding from Intel

- Significant possible performance gains – TEC, package cost
- Need to take advantage of temporal and spatial thermal landscape
- Need to be dynamically adaptable
- Need to be stable
- Difficult to do electro-thermal simulations
- Need thermal models for co-simulation
- Difficult to validate
- Best with “free-running core” (asynchronous, GALS, etc.)

Overview

- Motivation
- Figures of Merit (FOM)
- Temperature-Adaptive Circuits (TAC)
- **Temperature-Aware Computing (HotSpot)**
- Conclusions and future work

Hot spots in Power4

Temperature "landscape": space and time
How to estimate early in the design cycle?

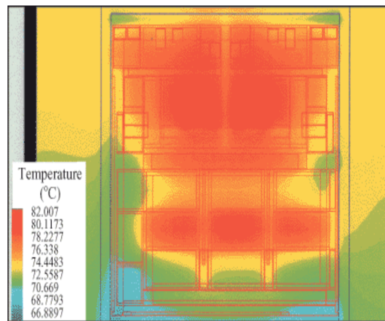


Figure 19

Map of FET junction temperatures for a 115-W packaged POWER4 chip derived from the chip power analysis and thermal modeling simulations described in the section on distribution.

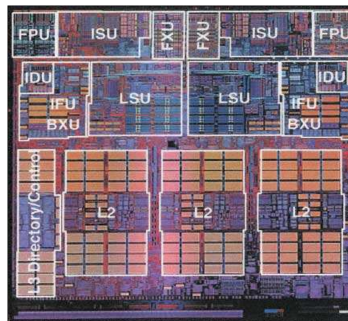
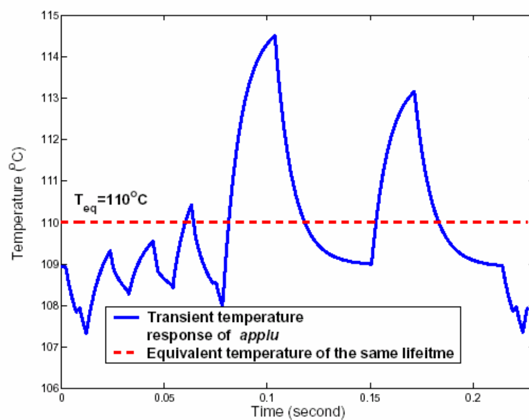


Figure 2

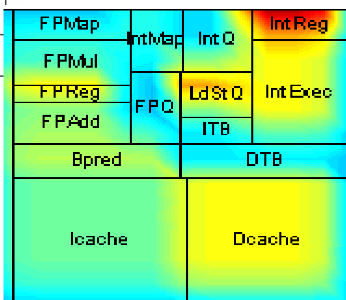
POWER4 chip photograph showing the principal functional units in the microprocessor core and in the memory subsystem.

Temporal, Spatial Variations



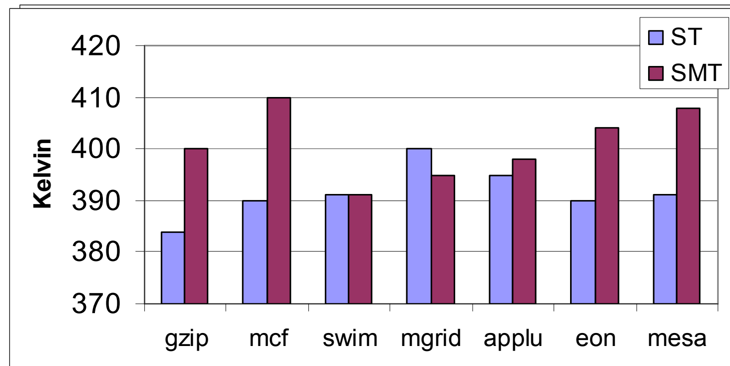
Temperature variation of SPEC applu over time

Hot spots increase cooling costs
⇨ must cool for hot spot



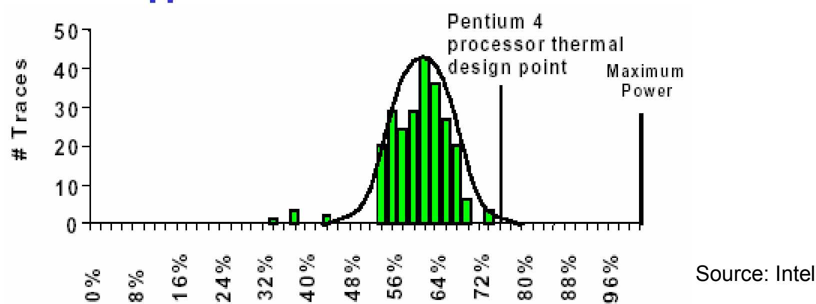
Application Variations

- **Wide variation across applications**
- **Architectural and technology trends are making it worse, e.g. *simultaneous multithreading* (SMT)**
  - **Leakage is an especially severe problem: *exponentially dependent on temperature!***



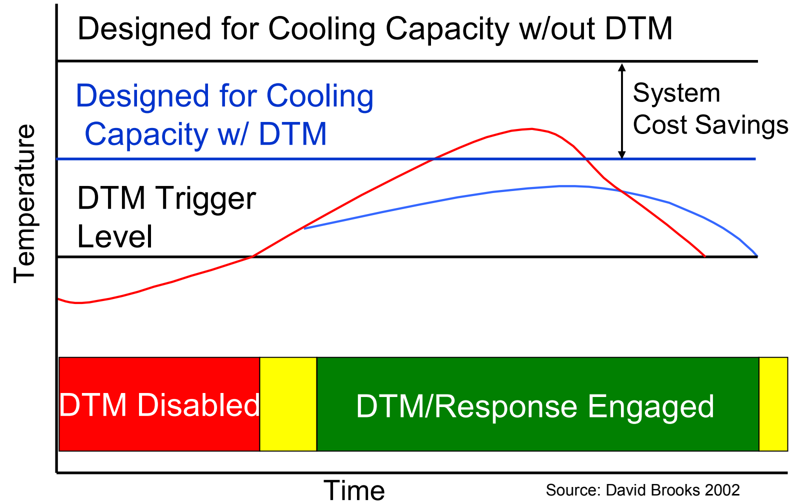
Pentium 4 Observations

- **For 200 traces (TPC-C, SPEC, Microsoft)**
  - **Thermal design point can be reduced to 75% of true "max power" with minimal performance loss**
    - **Aggressive clock gating**
    - **Application variations**



Source: Intel

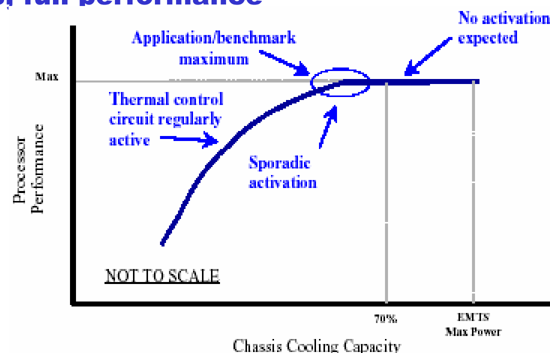
Dynamic Thermal Management



Brooks and Martonosi, HPCA 2001

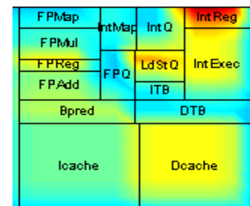
DTM

- **Worst case design is wasteful**
  - **Yet safe temperatures must be maintained**
- **Thermal monitors allow tradeoff**
  - **Cheaper package**
    - **More triggers, less performance**
  - **Expensive package**
    - **No triggers, full performance**



HotSpot project

- Collaboration between HPLP and LAVA Labs (ECE and CS depts. UVa)
- Deal with "hot spots"
  - Localized heating occurs much faster than chip-wide
    - microsec. to millisec.
  - Chip-wide treatment is too conservative
    - seconds to minutes
    - but there is significant lateral thermal coupling through the package
- How do we model this?



Thermal Modeling

- Performance simulator: e.g. SimpleScalar
- Power estimator: e.g. Wattch
- Want a fine-grained model of *temperature*
- Power dissipation
  - Will not capture hot spots
  - Too indirect
  - Does not account for power density

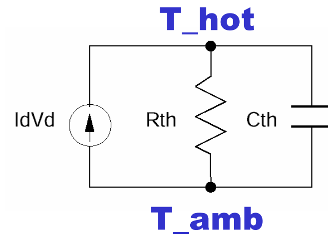
➔ **HotSpot** - a new model for localized temperature
  - Computationally efficient for use with power/performance simulators
  - Also extended for circuit level - grid

Dynamic Compact Thermal Model

Electrical-thermal duality

- V ≅ temp (T)
- I ≅ power (P)
- R ≅ thermal resistance (Rth)
- C ≅ thermal capacitance (Cth)
- RC ≅ time constant

T_hot



T_amb

KCL:

differential eq. $I = C \cdot dV/dt + V/R$

difference eq. $\Delta V = I/C \cdot \Delta t + V/RC \cdot \Delta t$

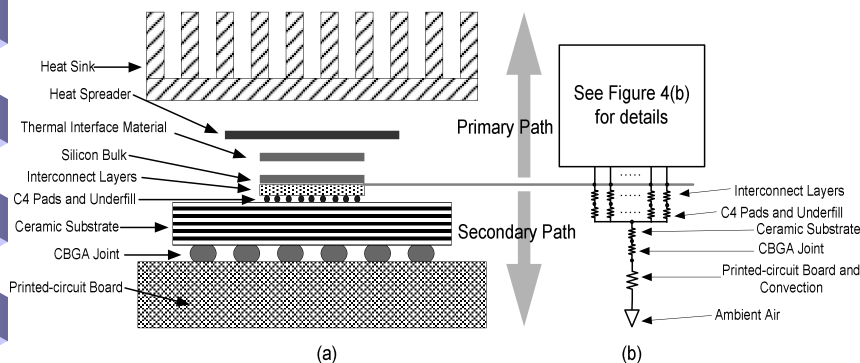
thermal domain $\Delta T = P/C \cdot \Delta t + T/RC \cdot \Delta t$

(T = T_hot – T_amb)

- One can compute stepwise changes in temperature for any granularity at which one can get P, R, C
- RC network ⇨ matrix form of these equations

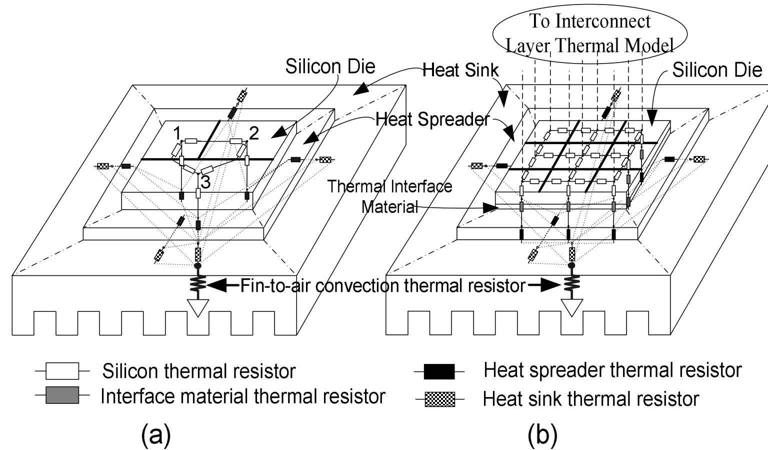
Model (package)

"Vertical" heat flow



Model (die)

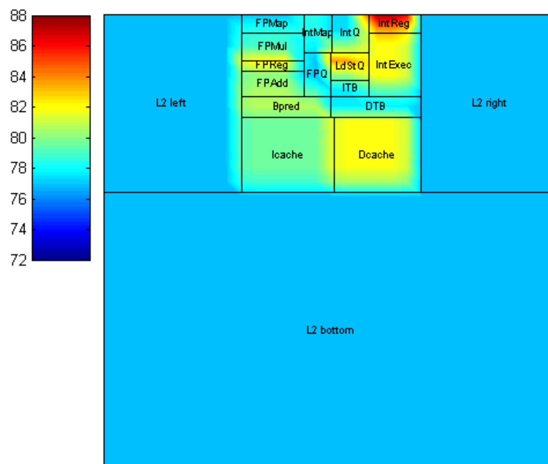
- Block granularity (architecture)
- Grid (circuits)
- Also lateral flow



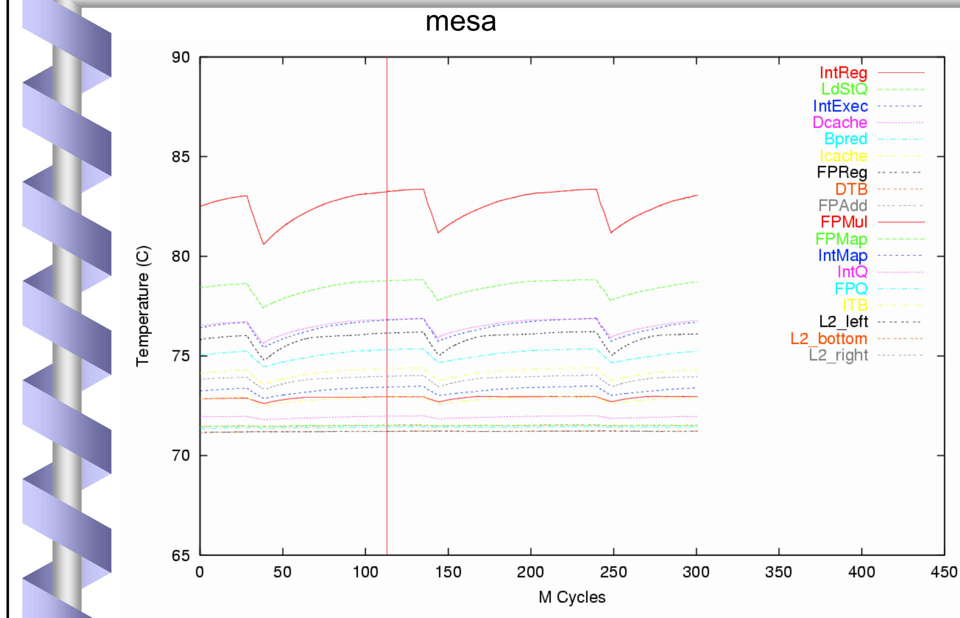
Thermal Modeling - Hot Spots

Deal with “hot spots”

– Either by throttling with DTM or by design



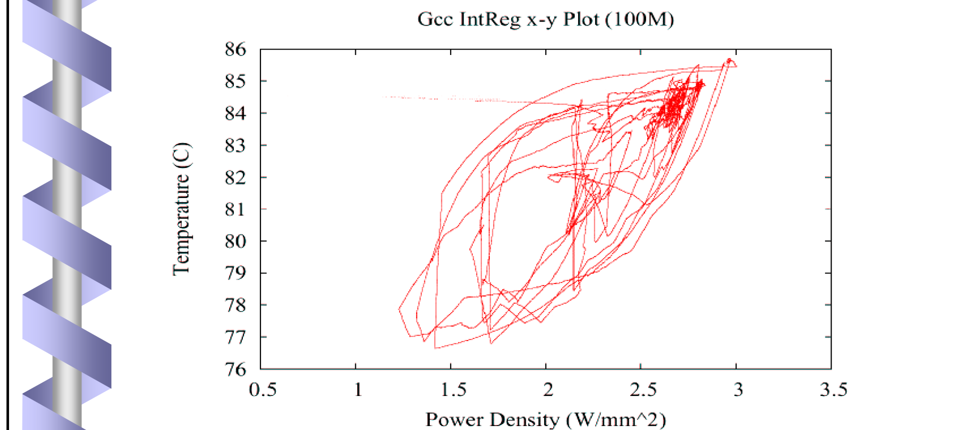
Time-Varying Behavior – Hot Spots



Thermal Modeling: P vs. T

Power metrics are not acceptable proxy

- **Chip-wide average will not capture hot spots**
- **Localized average will not capture lateral coupling**
- **Different units have different power densities**

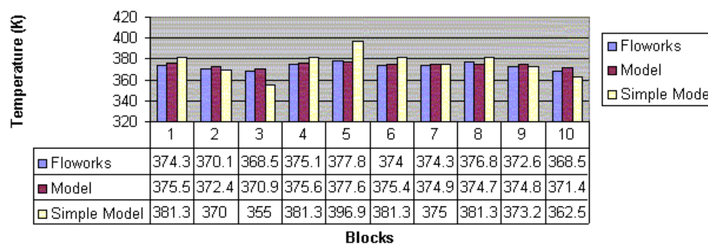


HotSpot validation: simulation

- **3% error relative to Floworks**
- **Lateral heat flow important**

Power for each block in Layout 3	Power1	Power 1 density (W/mm²)
Block1	10W	1.25
Block2	8W	0.8
Block3	2W	0.2
Block4	15W	1.25
Block5	15W	1.875
Block6	5W	1.25
Block7	12W	1
Block8	15W	1.25
Block9	13W	0.93
Block10	5W	0.5

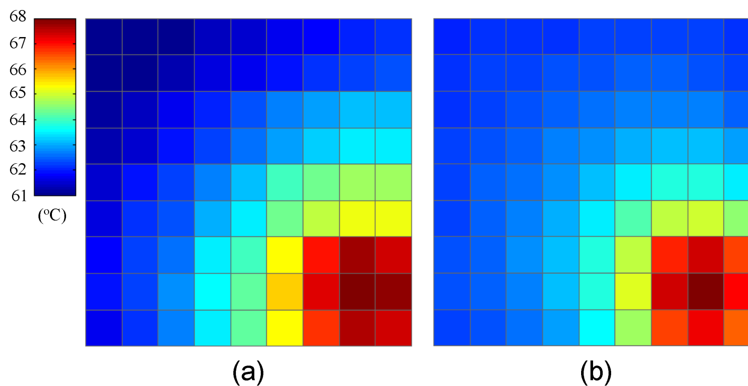
Layout3 Power Distribution1



	1	2	3	4	5	6	7	8	9	10
Floworks	374.3	370.1	368.5	375.1	377.8	374	374.3	376.8	372.6	368.5
Model	375.5	372.4	370.9	375.6	377.6	375.4	374.9	374.7	374.8	371.4
Simple Model	381.3	370	355	381.3	396.9	381.3	375	381.3	373.2	362.5

HotSpot validation: measurement

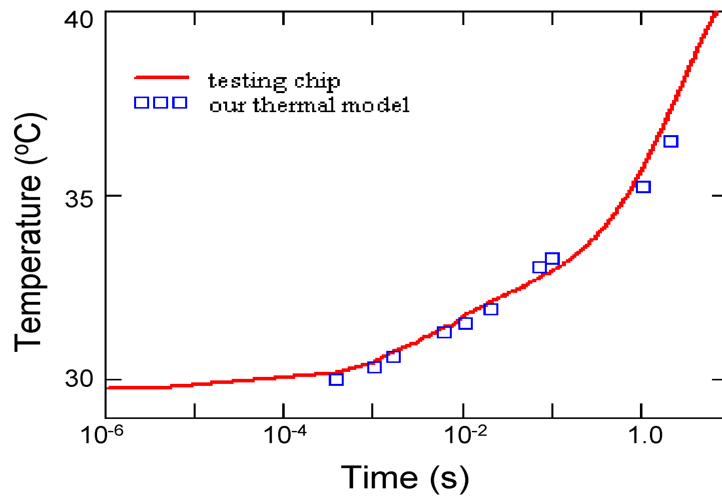
Micred test chip, steady state vs. HotSpot



(a) (b)

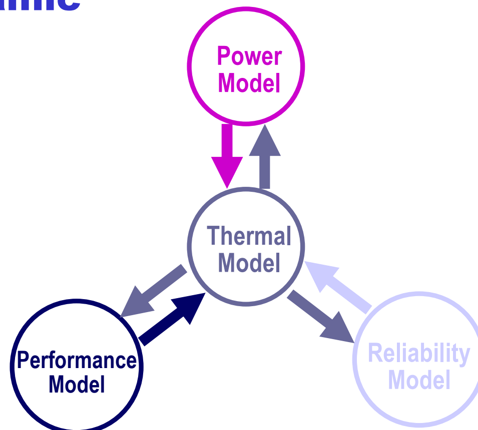
Validation: measurement

Micred test chip, transient vs. HotSpot

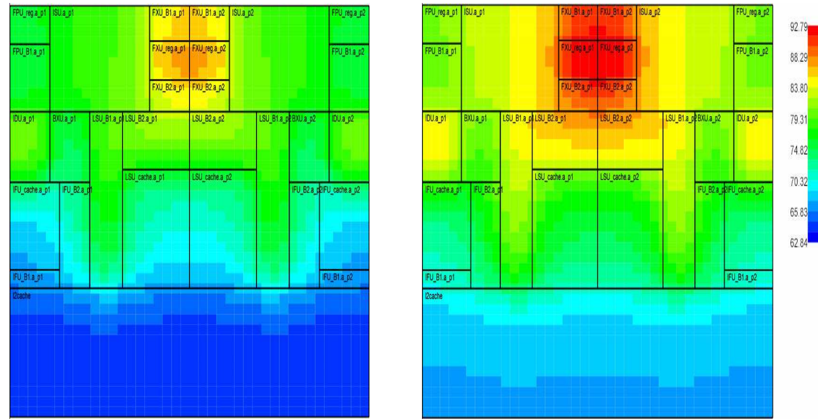


The Role of a Thermal Model

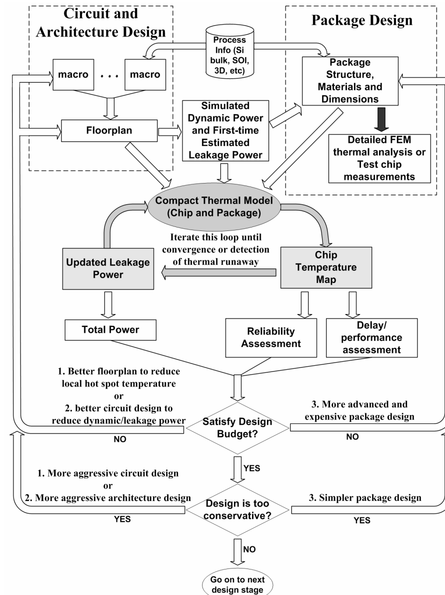
- helps close loop for accurate design estimations: static or dynamic



Self-consistent leakage



Design flow: work in progress!



HotSpot Summary

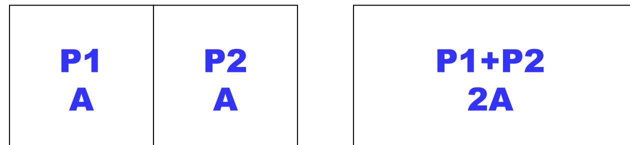
- HotSpot: simple, accurate and fast architecture and circuit level thermal model for microprocessors
- Models both steady-state and transient thermal behavior
- Per-functional unit or grid-based
- Specification of arbitrary floorplans
- Takes a power trace file as an input and outputs corresponding temperature trace
- Ability to modify package specifactions (type of interface material, size and type of heat spreader and heat sink etc.)
- Best paper award at ISCA 2003

Overview

- Motivation
- Figures of Merit (FOM)
- Temperature-Adaptive Circuits (TAC)
- Temperature-Aware Computing (HotSpot)
- Conclusions and future work

Temperature-aware FOM

- **Steady state: T~Power density (H), not just power (P)**
- **Spatial/temporal gradients: bad**
- **T-aware FOM = T^2 AD~H^2 AD = E^2/AD**
- **Truly 3-D, just thermals**



H1 = P1/A, H2 = P2/A, H = (P1+P2)/2A
TFOM1/AD = P1^2 TFOM2/AD = P2^2 TFOM/AD = (P1+P2)^2/2
P1^2 + P2^2 > 2P1P2 unless P1 = P2 (spatial gradients)
Similar for temporal gradients

Conclusions and future work

- **Temperature-adaptive circuits can increase performance or reduce power**
- **Thermal modeling allows early exploration of design space including thermal effects**
- **Variations will become prominent—shift from Deterministic to Adaptive design (for correlated chip-to-chip variations) and Probabilistic design (for random mismatch)**
- **Collaboration between academia and industry can lead to interesting opportunities**
- **HotSpot: >175 downloads since June'03**
  http://www.ece.virginia.edu/hplp
  http://lava.cs.virginia.edu/HotSpot
- **Still many challenges: future work!**

- Environment-aware sensor networks
- Power/thermal issues in mobile
- Power issues in graphics hardware
- Power issues in wireless sensor networks
- Integrated instruction of architecture and VLSI for CS/CE students
- Nanocircuits modeling, design and simulation
- Very-low voltage circuit design: eliminate leakage paths by using high fan-in pass-xtor logic

Looking for opportunities to collaborate: mircea@virginia.edu