Hybrid Dynamic Energy and Thermal Management in Heterogeneous Embedded MPSoCs

Shervin Sharifi*, Ayse K. Coskun+, Tajana S. Rosing*

*CSE Department, UC San Diego

+ECE Department, Boston University

(presented by Gunar Schirner, ECE Northeastern University)

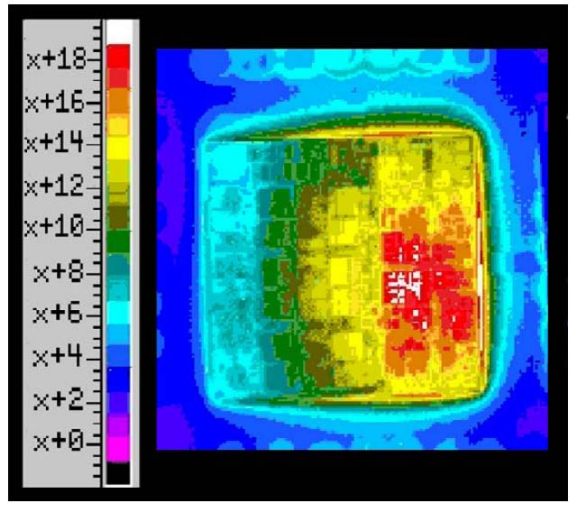
System Energy Efficiency Lab

Heterogeneous Multi-Processor SoCs

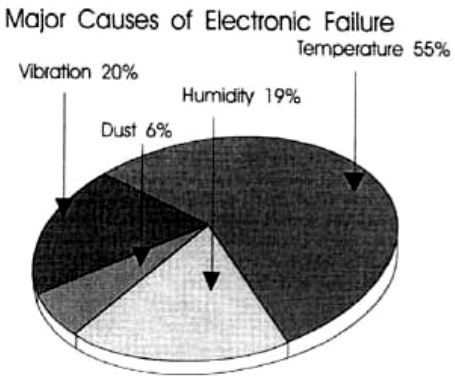
- Cores with various power and performance characteristics integrated on the same die
- Examples
  - Cell phones with ARM processors and DSPs
  - Wireless base stations with various types of processors and DSPs
  - SandBridge SB3500 SDR platform with 3 DSPs and one ARM9 core
- Challenges
  - Matching the resource needs of a thread with the available resources is challenging
  - Imbalance in power and temperature

Temperature Induced Problems

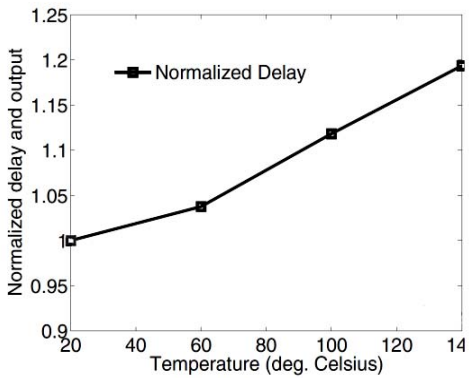
- Thermal hot spots
  - Accelerates failure mechanisms
  - Performance loss
  - Higher leakage power
- Temperature variations
  - Performance mismatch
  - Clock skew
  - Reliability issues
- Heterogeneity aggravates the thermal issues
  - Inherent disparity in power densities



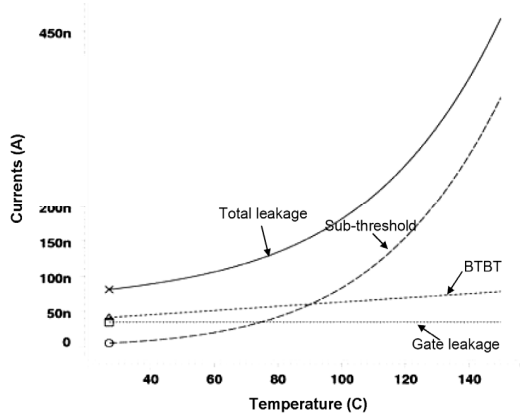
POWER5 processor, Jacobson, et. al., HPCA-11 2005



www.aitechnology.com



3 Nor gate stage ring oscillator
Bansal, et al. ASP-DAC 2006



Meterelliyoz, et al, ITC 2005

Related work

- [Lim, ISQED02]
  - A thermally-aware superscalar microprocessor with a secondary simpler lower power pipeline to migrate to at thermal emergencies
- [Heo, ISLPED03]
  - Power density is reduced by distributing the heat generation by migrating to spare units in cold areas of the chip

MicroArch level

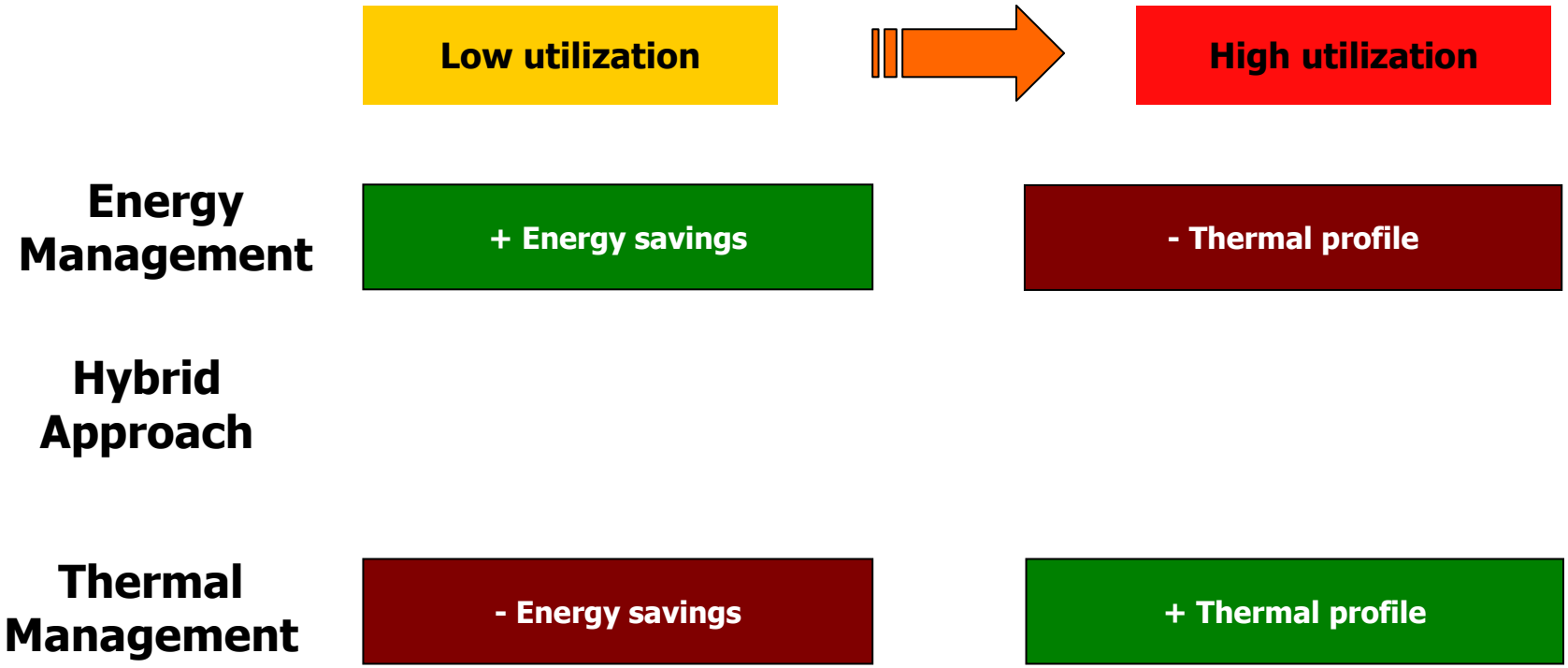
- [Kumar, Micro 03]
  - Proposed use of heterogeneous MPSoCs to achieve power-efficiency by migrating a process to the core that is expected to provide the best power-efficiency
- [Ghiasi, WCED04]
  - In a dual core processor, offloads workload to an extra low-power core to reduce the occurrence of thermal emergencies

System on Chip level

We focus on hotspots and thermal variations while managing energy in heterogeneous embedded MPSoCs where the types of the tasks run on the system are known a priori

C.-H., Lim, et al. “A thermal-aware superscalar microprocessor.” ISQED, 2002.
Heo, et al. "Reducing Power Density through Activity Migration“ ISLPED 2003.
Rakesh Kumar, et. al., “Single-ISA Heterogeneous Multi-Core Architectures: The Potential for Processor Power Reduction” Micro 2003
S. Ghiasi and D. Grunwald, “Design Choices for Thermal Control in Dual-Core Processors” Workshop on Complexity-Effective Design, 2004.

Energy and thermal management for heterogeneous embedded MPSoCs



Technique Overview

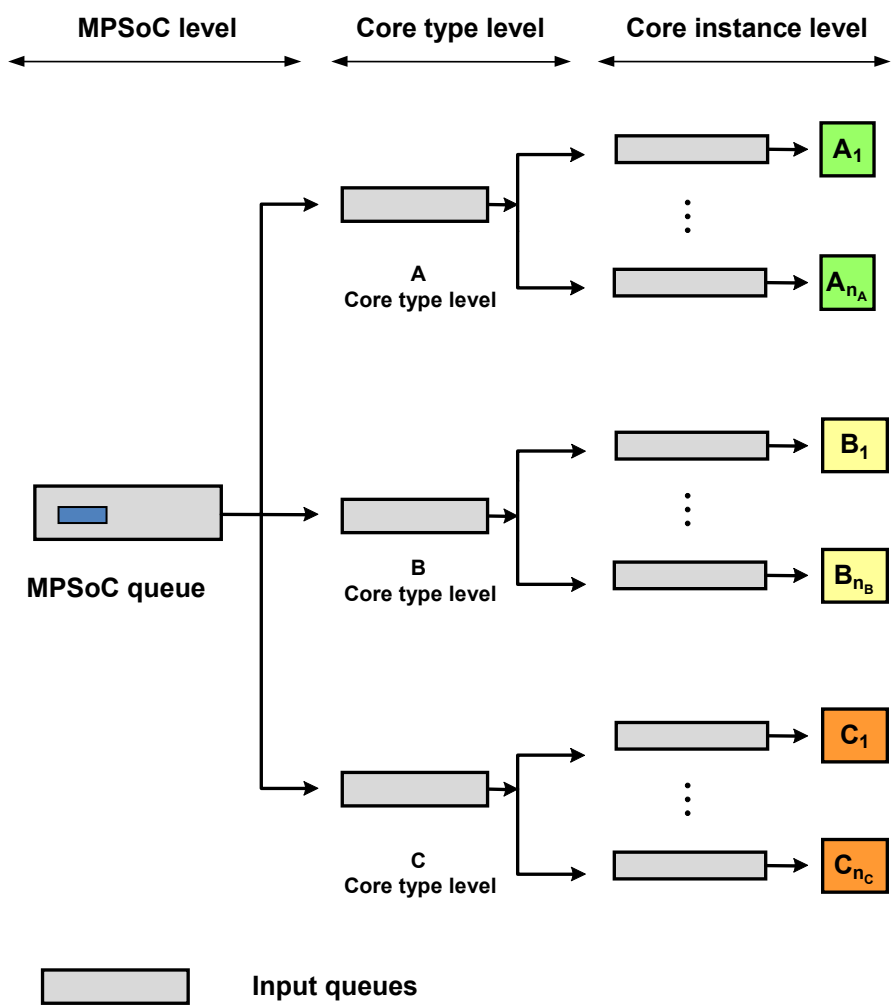
- Runtime measurement of workload demand
- Adjustment of processing capability of MPSoC in an energy and thermally aware fashion

- Supported by offline phases:
  - Task characterization
  - Voltage/Frequency calculation for thermal balancing

- Online modes:
  - Normal operation
    - Low and moderate utilizations
    - Energy saving by disabling cores
  - Thermal balancing
    - High utilization where thermal issues more likely to happen
    - Thermal management has a high priority

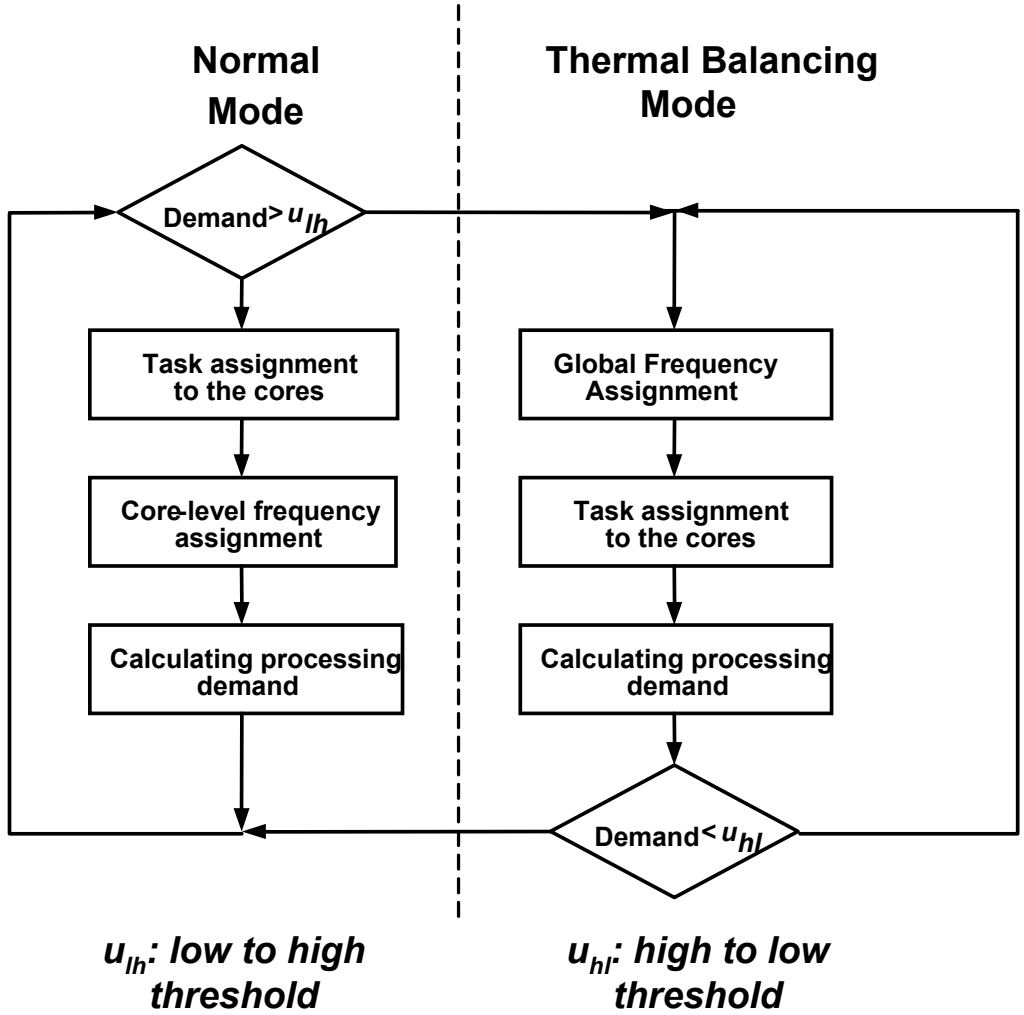
Task distribution hierarchy

- Hierarchical queue distribution
  - Allows modular control
  - Decouples the decisions between levels
- Three levels of queues
  - MPSoC level
  - Core type level
  - Core instance level
- Scheduling depending on online mode
  - Normal
  - Thermal balancing



Online modes

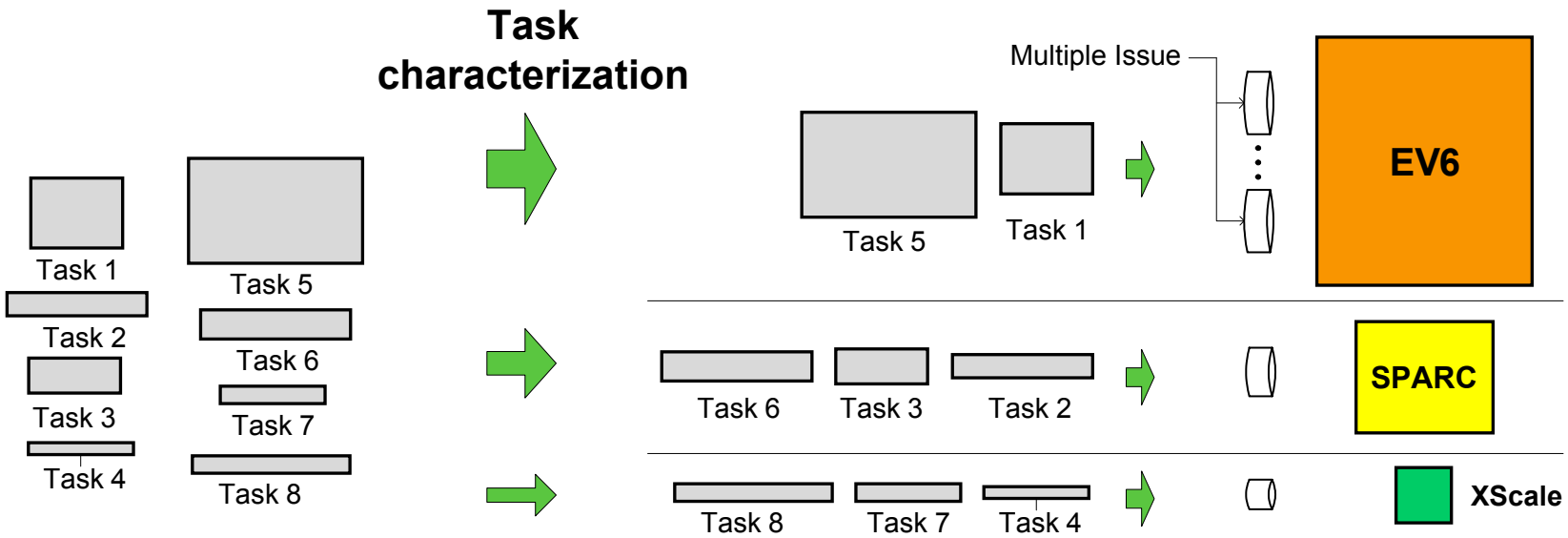
- Dynamic scheduler run at every clock tick
- Switch between modes based on demand
  - Queue content / status
- Normal mode
  - Task assignment
  - Core local frequency selection
- Thermal Balancing
  - Global VF
  - Task assignment to match targeted queue length



Offline phase

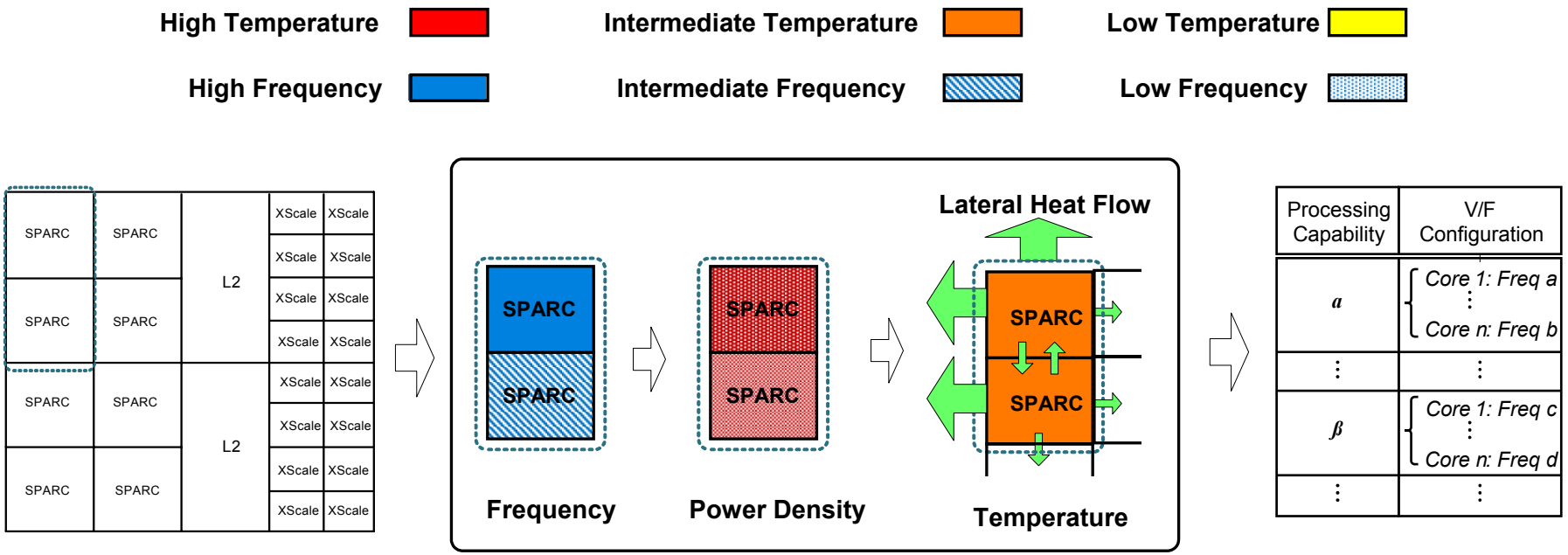
Task characterization

- Execute each task on each core with each VF setting
  - Record performance and power characteristics
    - E.g. Runtime, IPC, power consumption, #instructions
  - Determine suitable core(s) for execution



Thermal balancing

- Calculate DVFS settings for each core to balance temperature across chip at different utilization levels
  - Using HotSpot, thermal RC network



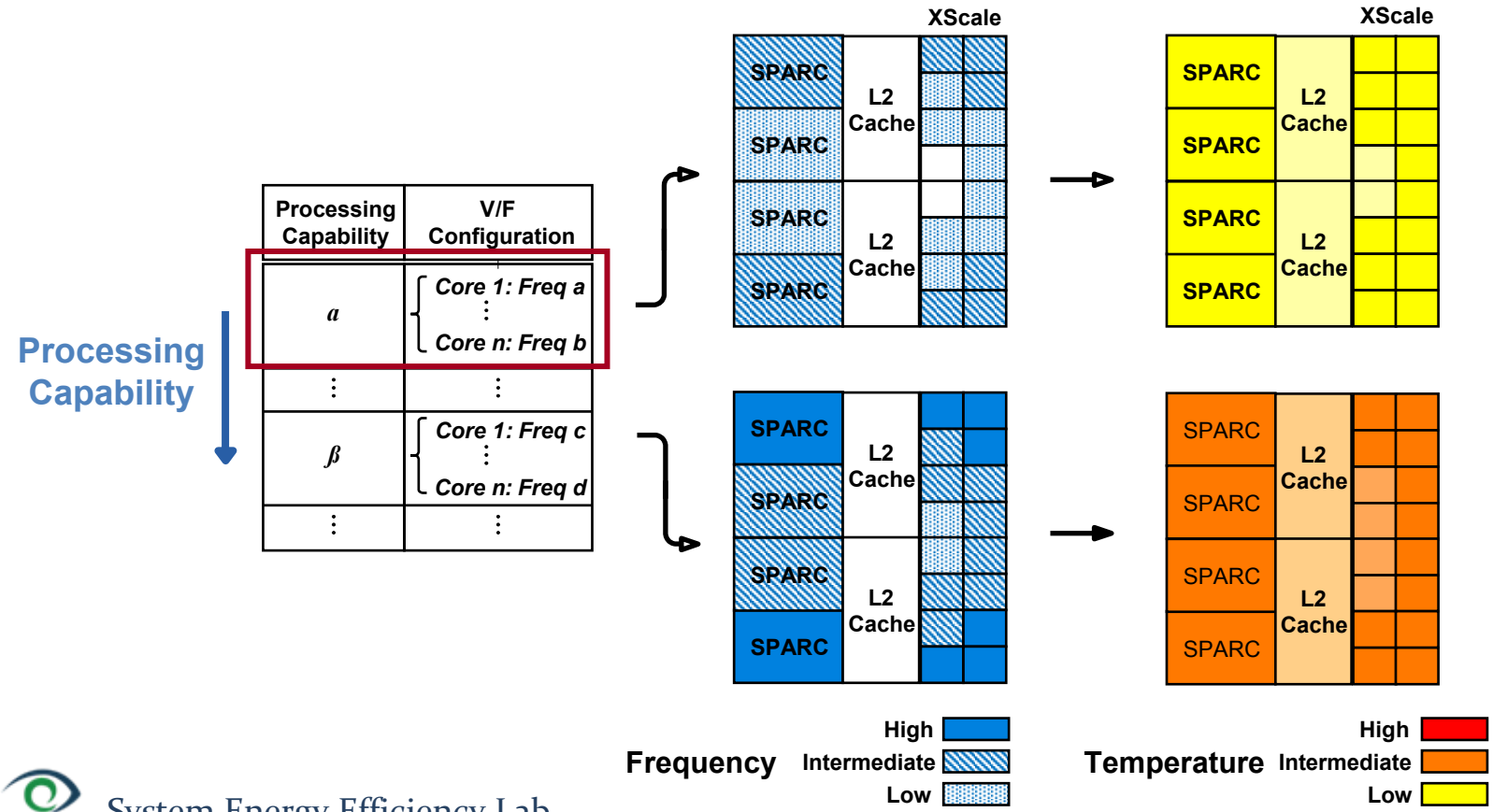
Online phase

Normal mode

- Calculate each task's arrival rate based on the length of the queues and the current service rate
  - Estimate queue length for each available frequency
- Queue length as a measure of workload demand
  - Continuously shrinking queue length: over-provisioning
  - Continuously growing queue length: under-provisioning
- Select DVFS setting to keep queue length constant
  - Continuously shrinking queues even at the lowest frequency
    - The core is turned off
  - Continuously growing queues even at the highest frequency
    - A new core is turned on

Thermal balancing mode

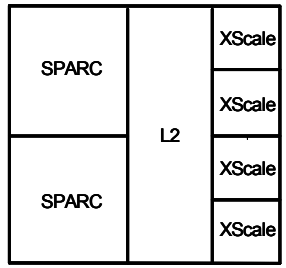
- Processing demand of the workload is calculated
- Pre-computed thermal balancing frequencies are assigned to the cores
- Core queues are balanced



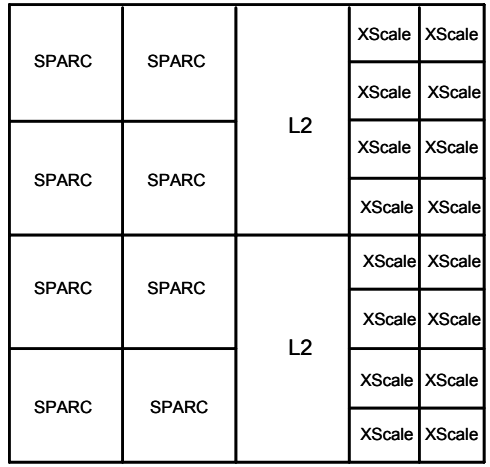
Experimental setup

- Two different types of in-order cores
  - XScale-like architecture
  - SPARC-like architecture
- Two different MPSoCs
- Tasks from MiBench benchmark suite
  - Varying load 45% - 95%
- M5 simulator integrated with Wattch
  - Performance and power
- HotSpot for thermal simulations

MPSoC1



MPSoC2



HotSpot Parameters

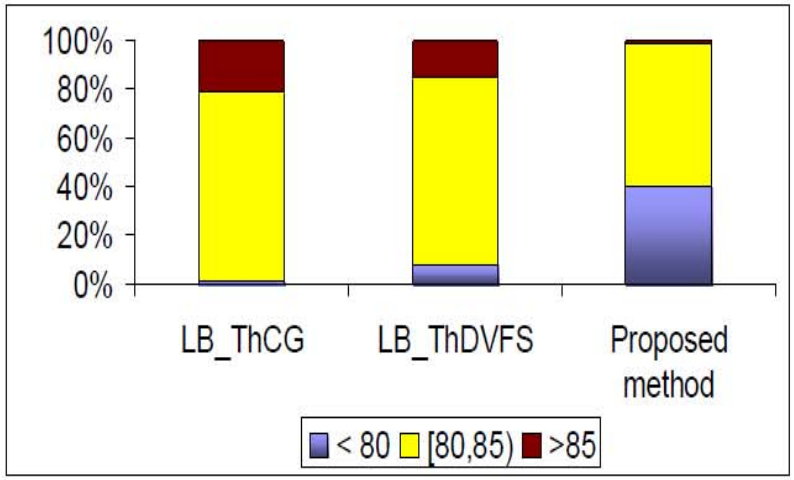
Die Thickness	0.15mm
Convection Capacitance	140J/K
Convection Resistance	1.5 K/W – 4 K/W

Baseline comparison techniques

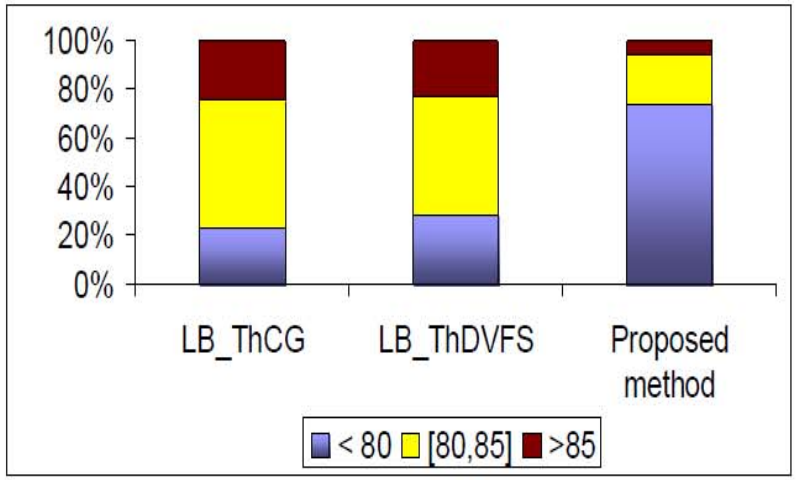
- Load balancing (LB)
  - Balances the number of waiting tasks among the available cores
- LB with clock-gating (*LB_ThCG*)
  - Clock is gated when core's temperature exceeds the threshold
- LB with temperature-triggered DVFS (LB_ThDVFS)
  - Frequency is reduced when temperature exceeds the threshold
  - If temperature remains above the threshold even at the lowest DVFS setting, the core is turned off

Results: Hotspots

- **An order of magnitude reduction in hotspots**



(a) MPSoC1

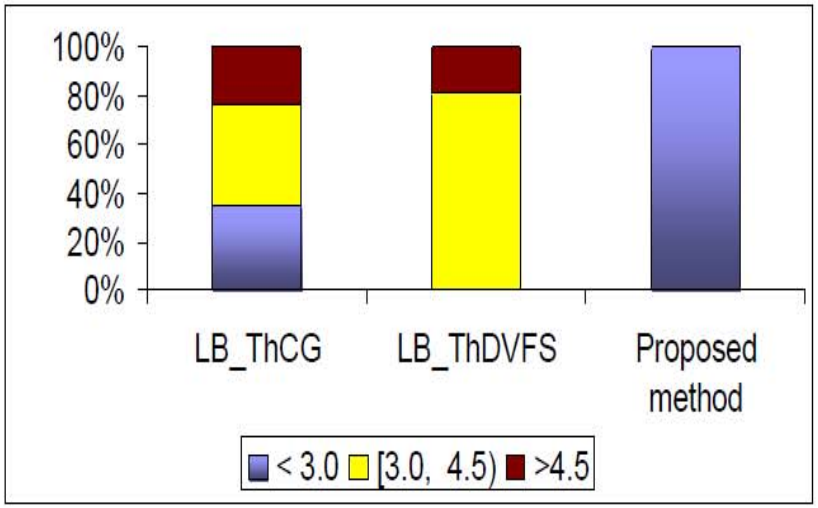


(b) MPSoC2

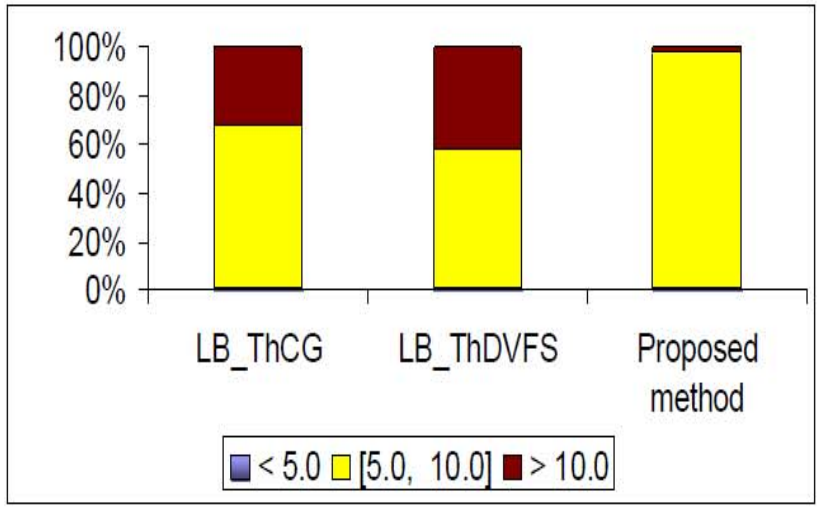
Distribution of hotspots HotSpots (°C)

Results: Temperature gradients

- **An order of magnitude reduction in high thermal gradients**



(a) MPSoC1

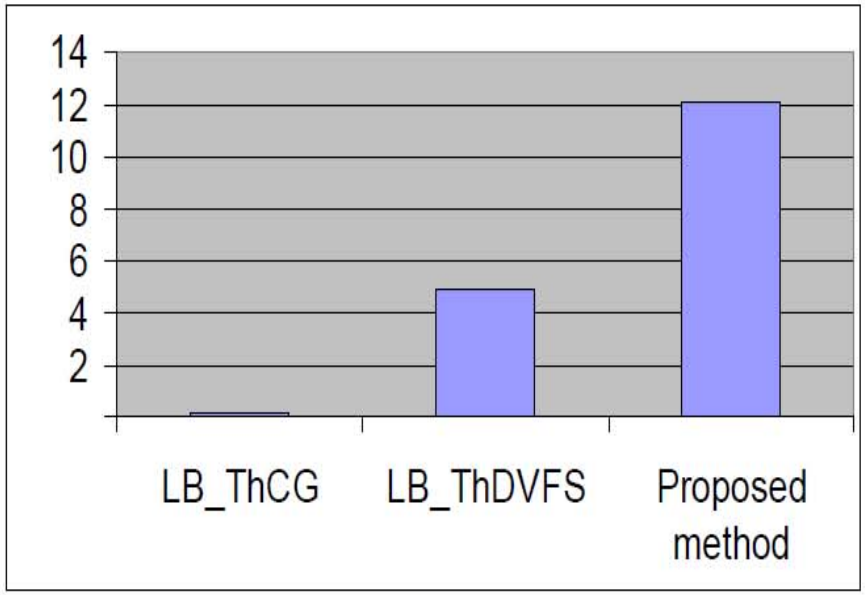


(b) MPSoC2

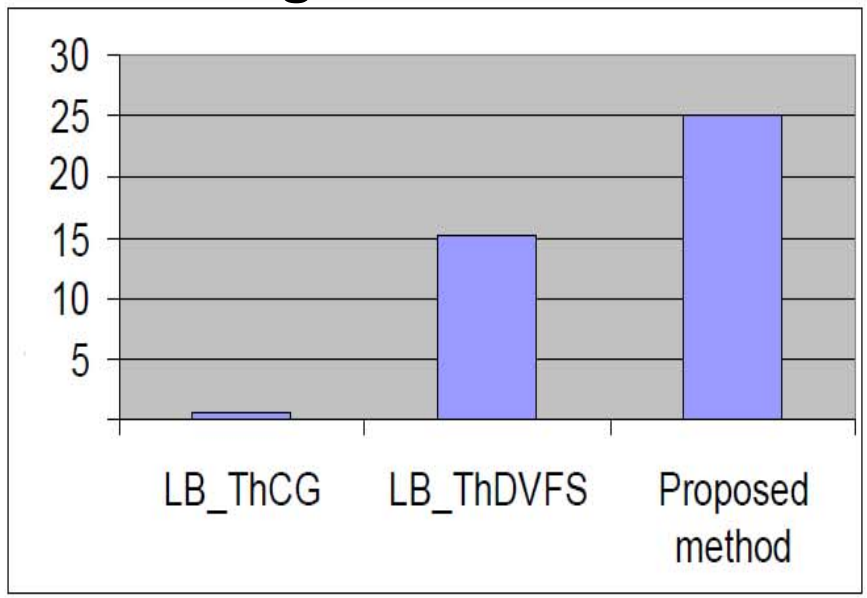
Spatial temperature variations (°C)

Results: Performance

- 12-25% improvement in average finish times of the tasks
  - Heterogeneity-aware task assignment
  - Lower thermal profile -> less stalling



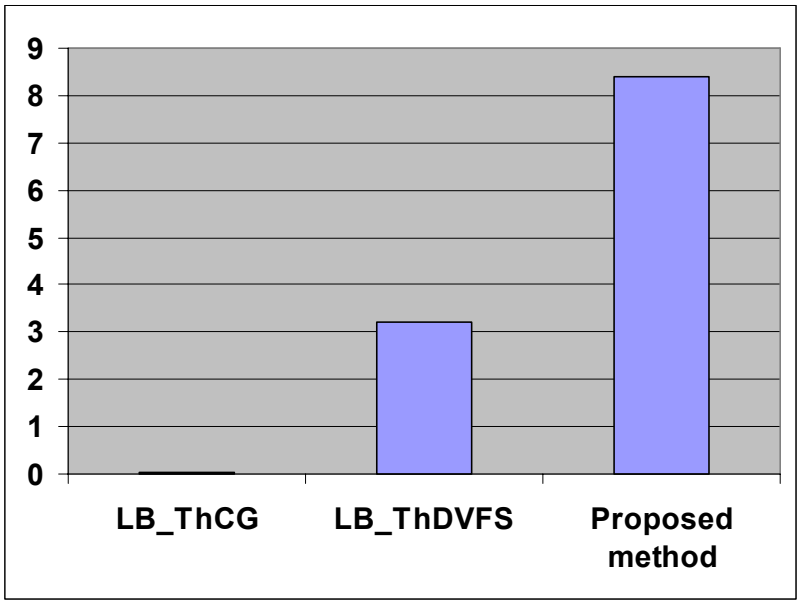
(a) MPSoC1

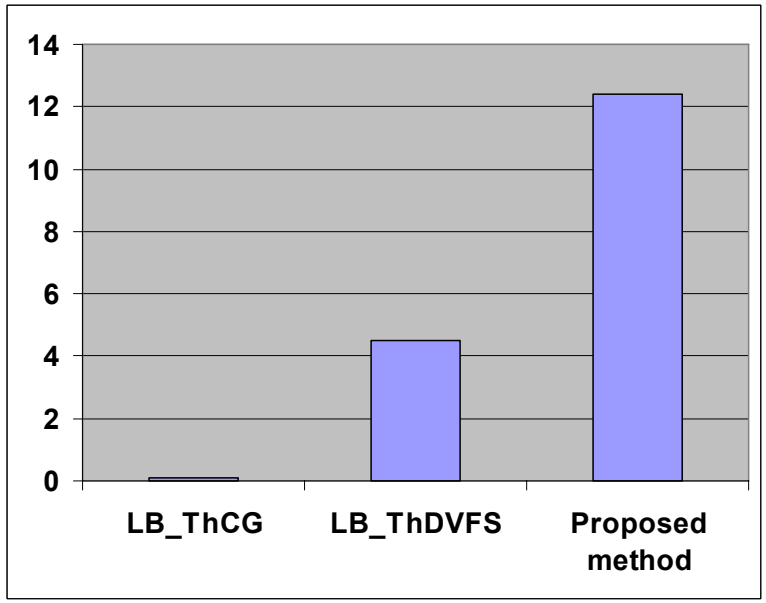


(b) MPSoC2

Results: Energy

- 8-12% energy improvement
  - Energy-aware task assignment in normal mode
  - Reduced hot spots -> lower leakage





Summary

- A dynamic thermal and energy management technique is proposed for heterogeneous embedded MPSoCs
- Hybrid approach
  - Normal mode
  - Thermal balancing mode
- Adjusts the processing capability to the workload demand
- Improves the energy savings by 12% and reduces thermal hotspots and spatial thermal gradients by an order of magnitude