

# Temperature-Constrained Power Control for Chip Multiprocessors with Online Model Estimation

**Yefu Wang**, Kai Ma, Xiaorui Wang

Department of EECS  
University of Tennessee, Knoxville



# Introduction

---

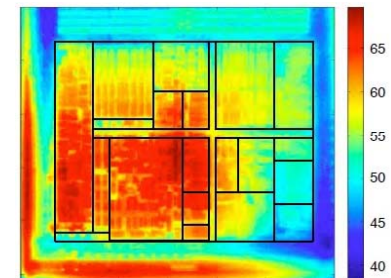
## ■ Power and temperature are serious concerns for processors

- Processors consume a major part of power in computer systems
- System failure due to overheating
- Cost of thermal packaging
- Increased level of core integration makes it worse



## ■ Power control for chip multiprocessors

- Peak power needs to be controlled
- Temperature must be kept lower than a threshold
- The performance delivered per watt needs to be maximized



[Mesa-Martinez'08]

# State of the Art

---

## ■ Power control for CMP

- Open-loop search or optimization [Isci'06], [Teodorescu'08], etc.
  - Highly dependent on the accuracy of the system model
- Heuristics [Isci'06], [Meng'08], etc.
  - No theoretical guarantee of control accuracy/stability
- Chip-wide DVFS (Dynamic Voltage and Frequency Scaling) [McGowen'06], [Floyd'07], etc.
  - Suboptimal in performance

## ■ Dynamic thermal management

- Heuristics or feedback control theory [Brooks'01], [Skadron'03], etc.
  - Power and temperature are controlled separately

## ■ Power/temperature management for server systems

- Server-level [Minerick'02], [Lefurgy'07], [Skadron'02], [Kephart'07], etc
- Server-rack-level [Kusic'16], [Wang'08], [Ranganathan'08], [Femal'05], etc
- Datacenter-level [Wang'09], [Fan'07], etc.

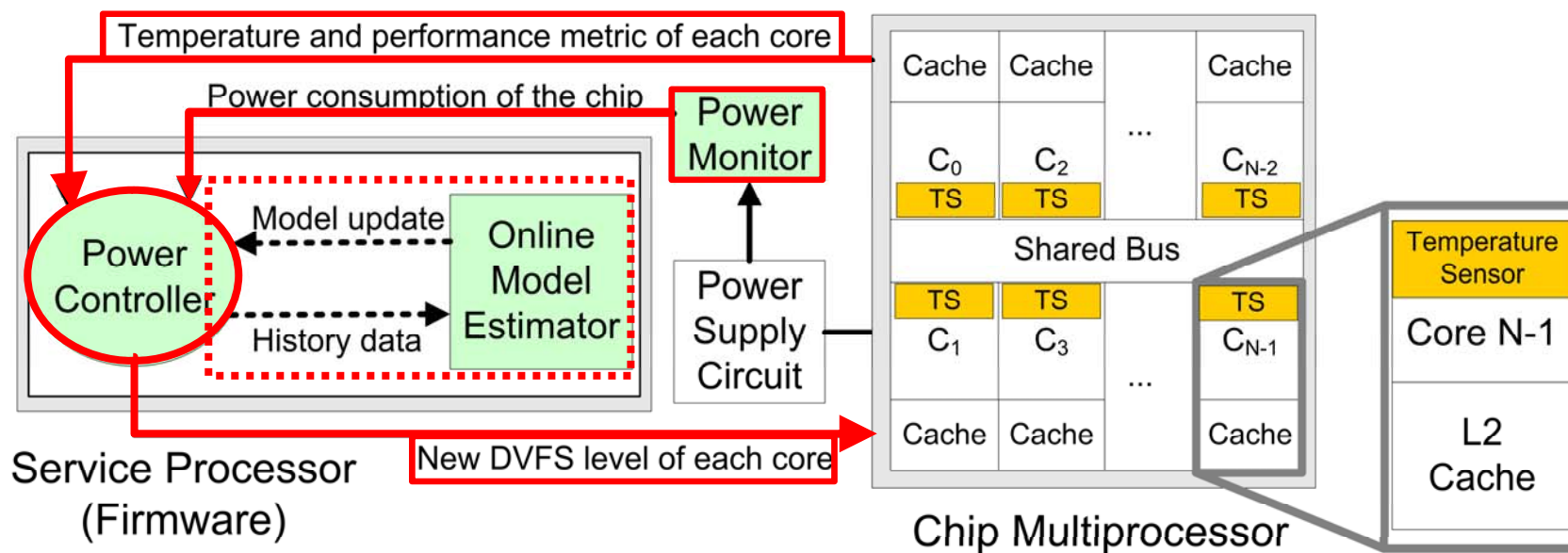
# Our Solution

---

- Actuator: per-core DVFS
- Manage power and temperature together with performance optimization
  - Power shifting among cores
  - Core variation and heterogeneity should be utilized to optimize processor performance
- Control-theoretic design
  - Multi-Input-Multi-Output (MIMO) control
    - To decide the DVFS levels of multiple cores
  - Model predictive control (MPC) theory
    - Well-established MIMO control theory with constraint
  - Theoretical guaranteed control performance and stability
- Online model estimation and correction
- Empirical results on hardware testbed

# Temperature-Constrained Power Control Loop

- MIMO control loop invoked periodically
  - Power monitor sends the chip-level power consumption to the controller
  - Controller reads temperature and performance metrics of each core
  - Controller computes new DVFS levels based on MPC control theory
  - New per-core DVFS levels are sent to the cores
  - Online model estimator updates the power model



# Steps of Model Predictive Control

---

## 1. System modeling

- Power model

## 2. Modeling the constraints

- Temperature constraint
- Physical frequency constraint
- Power budget

## 3. Controller design and analysis

- Problem formulation and solution
- Stability analysis

# Power Model

---

- Core level [Lefurgy'07], [Raghavendra'08]

$$p_i(k) = a_i f_i(k) + c_i \implies p_i(k+1) = p_i(k) + a_i \Delta f_i(k)$$

- Chip level

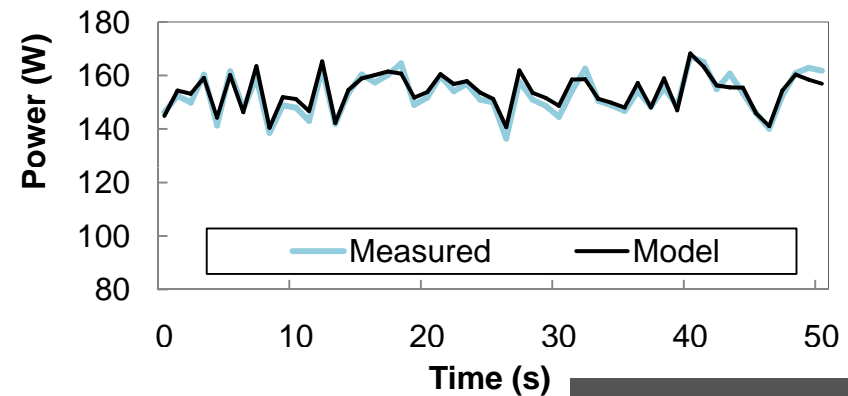
$$cp(k+1) = cp(k) + \begin{bmatrix} a_1 & \dots & a_N \end{bmatrix} \begin{bmatrix} \Delta f_1(k) \\ \vdots \\ \Delta f_N(k) \end{bmatrix}$$

Total Power

Frequency Vector

- Model estimation and validation

- System identification
- White noise input



# Temperature Model and Constraint

- From power to temperature [Han'07], [Brooks'07]

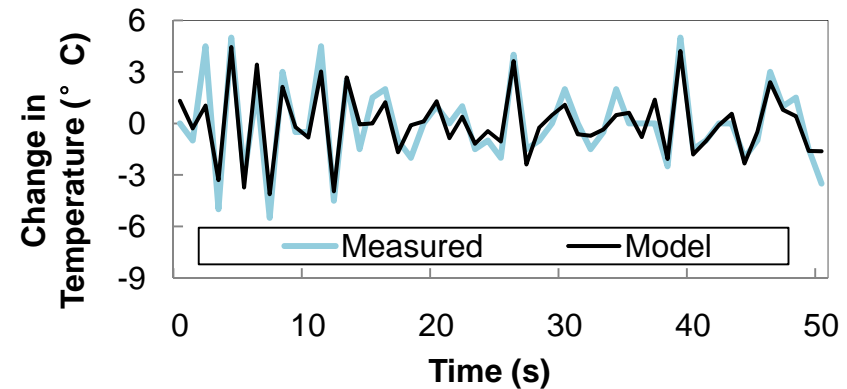
$$\mathbf{t}(\mathbf{k} + 1) = \mathbf{A}_T \mathbf{t}(\mathbf{k}) + \mathbf{B}_T \mathbf{p}(\mathbf{k})$$

- From frequency to temperature

$$p_i(k) = a_i f_i(k) + c_i \quad \Longrightarrow \quad \Delta \mathbf{t}(\mathbf{k}) = \mathbf{A}_T \Delta \mathbf{t}(\mathbf{k} - 1) + \mathbf{B} \Delta \mathbf{f}(\mathbf{k} - 1)$$

- Model estimation and validation

- White noise input
- System identification



- Temperature constraint

$$t_i < T_{i \max} \Rightarrow \mathbf{B}_i \cdot \mathbf{f}(\mathbf{k} + 1) < s_i$$

Constants



# Model Predictive Controller Design

---

Control objective: minimize the cost function

Model prediction

Reference trajectory

$$V(k) = \sum_{i=1}^P \underbrace{\|cp(k+i|k) - ref(k+i|k)\|_{Q(i)}^2}_{\text{Control Accuracy}} + \sum_{i=0}^{M-1} \underbrace{\|\Delta f(\mathbf{k} + \mathbf{i}|k) + f(\mathbf{k} + \mathbf{i}|k) - \mathbf{F}_{\max}\|_{\mathbf{R}(i)}^2}_{\text{Performance Optimization}}$$

Constraints:

- $F_{\min,j} \leq f_j(k+1) \leq F_{\max,j}$  → Physical frequency
- $cp(k) \leq P_s$  → Power budget
- $\mathbf{B}_i \mathbf{f}(\mathbf{k} + \mathbf{1}) < s_i$  → Overheating protection
- $f_i(k+1) = f_j(k+1)$  → Other design requirements

# Model Variation

---

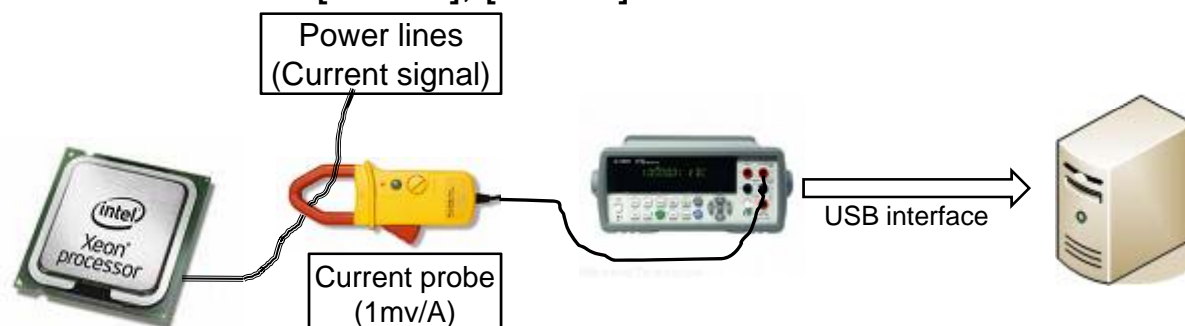
- Actual system model changes significantly at runtime
  - Unpredictable workload affects power behavior
  - Controller can be used on a different CMP
- Stability range
  - System is proven to be stable when the system model changes in a wide range
- Inaccurate model leads to degraded performance
  - Overshoot
  - Long settling time
- Use a standard recursive least square (RLS) estimator to correct the model periodically

# System Implementation

---

## ■ Testbed

- CPU: Xeon X5365
- Power monitor [Isci'03], [Wu'06]



- Temperature sensor: coretemp driver

## ■ Simulation environment

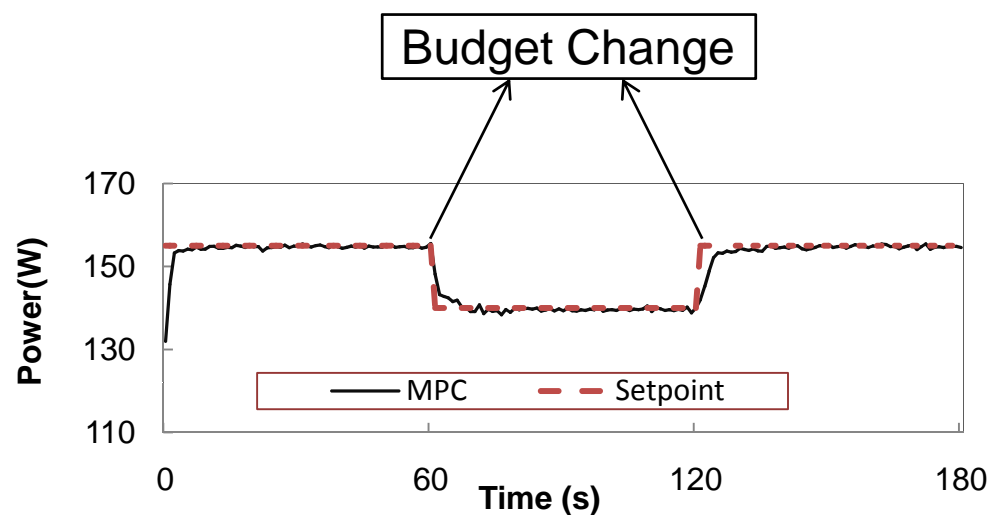
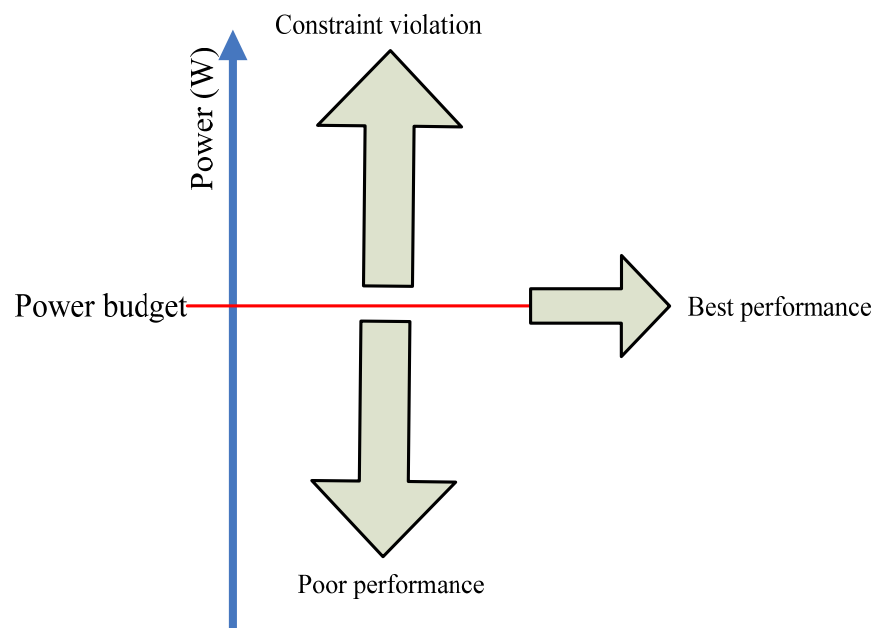
- CPU simulator: SESC with per-core DVFS support
- Power simulator: Wattch

## ■ Workload: SPEC CPU 2006

## ■ Core frequency modulator

- 4 discrete freq levels to approximate a fractional level?
  - For 2.89GHz, use 2.67, 3, 3, 2.67, 3, 3 ... on a smaller timescale (subintervals)

# Accurate Power Control



- MPC can precisely control the power of the CMP, with a standard deviation smaller than 1 W

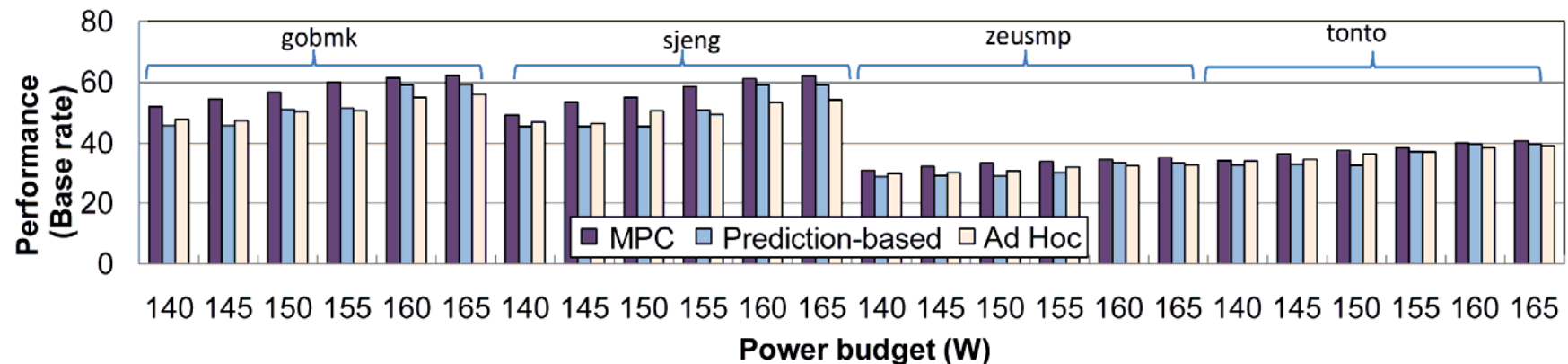
# Better Application Performance

## ■ Prediction-based

- Predict the power/performance of every DVFS combination based on an offline analysis.
- Select DVFS levels with the best performance under the power constraint

## ■ Ad Hoc: trial and error

- Power > budget: select a core and decrease its DVFS level by 1
- Power < budget: select a core and increase its DVFS level by 1



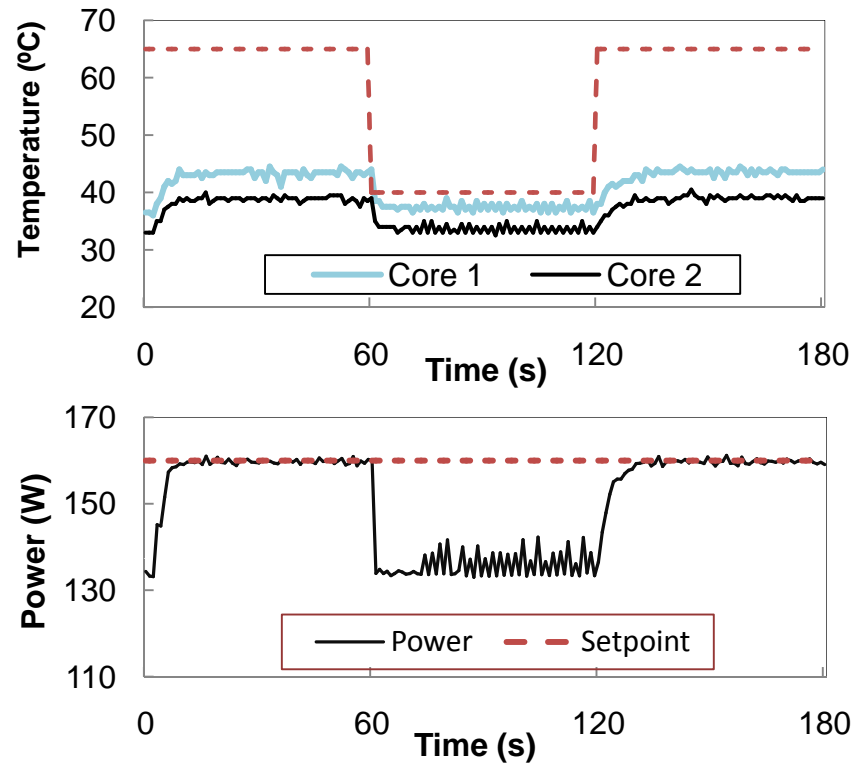
- More accurate power control and better application performance
- Confirmed by simulation results with 4, 8 and 16 cores

(More results available in paper)

# Temperature Constraint

---

- Emulate a thermal emergency by lowering the temperature constraint
  - Temperatures are quickly constrained to stay below the desired value
  - Power consumption reduced for temperature reduction



# Conclusion

---

- A temperature-constrained chip-level power controller
  - Designed based on MPC control theory
  - Accurately controls power consumption
  - Temperatures of the cores are limited to stay below the constraint.
  - An online model estimator periodically updates the system model
- Compared with state-of-the-art work
  - More accurate power control
  - Better application performance

# Acknowledgement

---

This work was supported by

- NSF CAREER Award CNS-0845390
- NSF CSR Grant CNS-0720663
- Power-Aware Computing Award, Microsoft Research
- Office of Naval Research (N00014-09-1-0750)