Temperature-Constrained Power Management Scheme for 3D MPSoC

Arnica Aggarwal, Sumeet S. Kumar, Amir Zjajo, Rene van Leuken

Delft University of Technology
The Netherlands.

14th May 2012
INTRODUCTION

• Better system integration and performance by scaling down device size
• Further scaling is becoming challenging due to limitations of:
  - interconnect performance
  - leakage power
  - process variation
• 3D integration offers increased system integration and performance
• Increased power density
### MOTIVATION

- Temperature constraints for power management is important
- DVFS is a widely used power management technique

<table>
<thead>
<tr>
<th>2D MPSoC</th>
<th>3D MPSoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Monitor PE workload <em>only</em></td>
<td></td>
</tr>
<tr>
<td>• PE temp. monitored <em>independently</em></td>
<td></td>
</tr>
</tbody>
</table>

*PE: Processing Element*
MOTIVATION

- Temperature constraints for power management is important
- DVFS is a widely used power management technique

<table>
<thead>
<tr>
<th>2D MPSoC</th>
<th>3D MPSoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Monitor PE workload <strong>only</strong></td>
<td></td>
</tr>
<tr>
<td>- PE temp. monitored <strong>independently</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Need for new approaches to deal with the differences**

**PE:** Processing Element
OUTLINE

- System modeling
- Power management algorithm
- Experimental setup
- Results
- Conclusions
SYSTEM MODELING (I)

New DVFS level for each PE

Chip Multi-processor

Power supply circuit

Power Management Block

Power monitor

Power consumption of the chip

Temperature and activity of each PE

TS: Temperature Sensor
PE: Processing Element
SYSTEM MODELING (II)

- Each PE has fixed V/F levels: VF1, VF2, VF3, VF4, VF5, VF6

- Relation between Power and DVFS level can be given by:
  \[ P = A \cdot (V^2F) + B \]

  \[ \Rightarrow \Delta P = A \cdot \Delta(V^2F) \]
Thermal Model Equations

\[ T_1 = T_{amb} + (x \cdot P_1 + P_2) \cdot R_{hs} \]
\[ T_2 = T_1 + (-y \cdot P_1 + P_2) \cdot R_{hs} \]

Direct relation between temperature and power

\[ \Delta T_i = R_{\text{eff}_{ij}} \cdot \Delta P_j \]

\[ R_{\text{eff}} = \begin{pmatrix} R_{11} & R_{12} & R_{13} & R_{14} \\ R_{21} & R_{22} & R_{23} & R_{24} \\ R_{31} & R_{32} & R_{33} & R_{34} \\ R_{41} & R_{42} & R_{43} & R_{44} \end{pmatrix} \]

\[ \Delta P = A \cdot \Delta (V^2F) \implies \Delta T = R_{\text{eff}} \cdot A \cdot \Delta (V^2F) \]
POWER MANAGEMENT ALGORITHM

New approach: Proposed approach

2D approach: approach used in 2D MPSoC
PE temp. monitored independently
POWER MANAGEMENT ALGORITHM

START

Wait for \( f \) on control cycle

Initial Updates

(Temp-check cycle = 1) ?

YES

Thermal Runout

Convergency check

(converge = true) ?

NO

Pull-up / Pull-down

YES

Write-back and reset

Control period: PMB decides new V-F levels
Temp-check Period: current temp. available to PMB
POWER MANAGEMENT ALGORITHM

START

Wait for f on control cycle

Initial Updates

(Temp-check cycle = 1) ?

YES

Thermal Runout

NO

Convergency check

(Converge = true) ?

NO

Pull-up / Pull-down

YES

Write-back and reset

Control period: PMB decides new V-F levels
Temp-check Period: current temp. available to PMB
POWER MANAGEMENT ALGORITHM

START

Wait for \( f \) on control cycle

Initial Updates

(Temp-check cycle = 1)?

YES

Thermal Runout

NO

Convergency check

(converge = true)?

NO

Pull-up / Pull-down

YES

Write-back and reset

Control period: PMB decides new V-F levels
Temp-check Period: current temp. available to PMB
**Thermal Runout**

- If temperature crosses safety limit (e.g. 328 K for critical temp of 330 K)
  - Operating V-F levels are adjusted

- 2D approach: V-F of victim PE is scaled down

- New approach: Weight of each PE = \( (a * (1\text{-utilization})) + (b * R_{eff}) \)

\( R_{eff} = \begin{pmatrix}
  R_{11} & R_{12} & R_{13} & R_{14} \\
  R_{21} & R_{22} & R_{23} & R_{24} \\
  R_{31} & R_{32} & R_{33} & R_{34} \\
  R_{41} & R_{42} & R_{43} & R_{44}
\end{pmatrix} \)

*Utilization*: PE activity in previous control period
POWER MANAGEMENT ALGO. (III)

START

Wait for $f$ on control cycle

Initial Updates

(Temp-check cycle = 1)?

YES: Thermal Runout

NO: Convergency check

(converge = true)?

NO: Pull-up / Pull-down

YES: Write-back and reset

**Power budget:** Constrains peak power of the chip
POWER MANAGEMENT ALGO. (III)

START

Wait for \( f \) on control cycle

Initial Updates

(Temp-check cycle = 1) ?

YES

Thermal Runout

NO

Convergency check

(\text{converge = true}) ?

NO

Pull-up / Pull-down

YES

Write-back and reset

\textbf{Power budget:} Constrains peak power of the chip
POWER MANAGEMENT ALGO. (IV)

*pull-up / pull-down*

- **2D approach:** V-F scaled according PE utilization **ONLY**

- **New approach:** V-F scaled according to a weighted equation
  
  \[(c \times \text{utilization}) + (d \times \text{temp margin}) + (e \times \text{height}) + (f \times \text{area})\]

- V-F of a PE is scaled up only if
  - \((T + \Delta T)\) does not cross safety limit
  - Was not scaled down in *Thermal Runout* stage
POWER MANAGEMENT ALGO. (V)

START

→ Wait for \( f \) on control cycle

→ Initial Updates

→ (Temp-check cycle = 1)?
  → YES
  → Thermal Runout
  → NO
  → Convergency check

→ (converge = true)?
  → NO
  → Pull-up / Pull-down
  → YES
  → Write-back and reset
EXPERIMENTAL SETUP (I)

**SimpleScalar Configured with Wattch** + **Application from a benchmark**

- Activity trace for PEs
  - PE 0
  - PE 1
  - PE 2

- Power values for PE
  - V/F levels
  - PE 0
  - PE 1
  - PE 2

**Linear Regression**

\[ P = A^*V^2 + B \]

**Before run-time** (data collection)

**TESTBENCH**

- Calculate: utilization (each PE), power (each PE), total power

**run-time** (simulation)

**Power Management Block**

**Input files for 3D-ICE**

- die 0
- die 1
- stack

**3D-ICE Thermal Simulator**

- Final temperature report
  - temperature of all PEs
EXPERIMENTAL SETUP (II)

- Control period: 50us (60,000 cycles at max freq)
- Temperature input: every 1ms
- Temperature constraint: 320 K on each PE (strict)
- Safety margin of: 2K

- 6 DVFS levels per PE are taken
  - Frequency: 700 – 1200MHz
  - Voltage: 0.8V – 1.1V

- Deep sleep mode: only clock is gated
- Basicmath application from MiBench Benchmark
- Two sets of simulation setup
RESULTS (I)

SUM OF FREQUENCIES OF ALL PEs

TOTAL POWER

Frequency (MHz)

Power (W)

Time (ms)

3D algo
2D algo

3D algo
2D algo

3D power budget
2D power budget
RESULTS (II)

TEMPERATURE OF PE 0

Temperature (K)

Time (ms)

- new approach
- 2D approach
- critical temp
- Safety limit
RESULTS (III)

- Similar losses on all layers
- Similar execution time
- Total execution time reduces

19.55% reduction in total execution time
CONCLUSIONS

- A new power management scheme was proposed
- High level thermal model was achieved
- Weighted equations were used for choosing V-F levels
- Power and temperature were maintained below constraints
- Less temperature fluctuations
- Up to 19.55% improvement in total execution time
  - High Sum of Frequencies
  - Turning OFF of PEs was avoided
Temperature-Constrained Power Management Scheme for 3D MPSoC

Arnica Aggarwal, Sumeet S. Kumar, Amir Zjajo, Rene van Leuken

Delft University of Technology
The Netherlands.

14th May 2012