Thermal-aware Post Compilation for VLIW Architectures

Wen-Wen Hsieh and TingTing Hwang

National Tsing Hua University, Taiwan
Outline

• Introduction
• Motivation
• The Proposed Methods
  – Binding Method
  – Forwarding Method
• Experimental Results
• Conclusions
Outline

• Introduction
• Motivation
• The Proposed Methods
  – Binding Method
  – Forwarding Method
• Experimental Results
• Conclusions
Introduction (1/2)

• Transistor density and power consumption have grown rapidly.
  → Serious heat dissipation problem
• High temperature induces undesirable effects.
  – Low reliability
  – Low performance
  – High cooling costs
Introduction (2/2)

• A thermal management method needs to be developed.
  – To reduce hotspots
  – To balance the temperature distribution
Outline

- Introduction
- **Motivation**
- The Proposed Methods
  - Binding Method
  - Forwarding Method
- Experimental Results
- Conclusions
Temperature Distribution

(°C)

[Bar chart showing temperature distribution for various components: DL1, IL1, L2, BTB, Decode, IALU, FALU, IREG, FREG, DTLB, ITLB. The IREG component has the highest temperature.]
Register File Architecture

enable signal

Conventional Register File

Sub-bank Register File

Proposed in [Tseng 2003]
Outline

• Introduction
• Motivation
• The Proposed Methods
  – Binding Method
  – Forwarding Method
• Experimental Results
• Conclusions
The Proposed Method

• A static thermal management technique at compiler level.
• Target at VLIW architecture.

• We propose two techniques.
  – Register binding
    → balance the temperature of register file
  – Forwarding method
    → reduce the access count of register file
Outline

• Introduction
• Motivation
• The Proposed Methods
  – Binding Method
  – Forwarding Method
• Experimental Results
• Conclusions
Register Binding
– Simple Binding

$U_{1_{\text{Temp}}} > U_{2_{\text{Temp}}} = U_{3_{\text{Temp}}} > U_{4_{\text{Temp}}}$

<table>
<thead>
<tr>
<th>U1</th>
<th>U2</th>
</tr>
</thead>
<tbody>
<tr>
<td>REG_1</td>
<td>REG_2</td>
</tr>
<tr>
<td>REG_3</td>
<td>REG_4</td>
</tr>
<tr>
<td>U3</td>
<td>U4</td>
</tr>
</tbody>
</table>

Binding order:

REG_1 → REG_1 → REG_1 → REG_1

Hotspot: REG_1
Register Binding
– Round-Robin-Like Binding

\[ U_1^{\text{Temp}} > U_2^{\text{Temp}} = U_3^{\text{Temp}} > U_4^{\text{Temp}} \]

<table>
<thead>
<tr>
<th>U1</th>
<th>U2</th>
</tr>
</thead>
<tbody>
<tr>
<td>REG_1</td>
<td>REG_2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Binding order:

\[ \text{REG}_1 \rightarrow \text{REG}_2 \rightarrow \text{REG}_3 \rightarrow \text{REG}_4 \]

Hotspot: \( \text{REG}_1 \) & \( \text{REG}_2 \)
Register Binding
– Floorplan-Aware Binding

\[ U_{1_{\text{Temp}}} > U_{2_{\text{Temp}}} = U_{3_{\text{Temp}}} > U_{4_{\text{Temp}}} \]

<table>
<thead>
<tr>
<th>U1</th>
<th>U2</th>
</tr>
</thead>
<tbody>
<tr>
<td>REG_1</td>
<td>REG_2</td>
</tr>
<tr>
<td>REG_3</td>
<td>REG_4</td>
</tr>
<tr>
<td>U3</td>
<td>U4</td>
</tr>
</tbody>
</table>

Binding order:

\[ \text{REG}_4 \rightarrow \text{REG}_2 \rightarrow \text{REG}_3 \rightarrow \text{REG}_1 \]

Temperature Balance !!
Register Binding

- **Spatial information**
  - Floorplan
    - location of every unit

- **Temporal information**
  - Control flow graph (CFG)
    - program flow
  - Temperature trace file
    - temperature of every unit
Penalty Function

- Determine whether a sub-bank register file is selected for binding or not.
- Apply *Penalty* to every sub-bank register.
- Bind the target live range to the bank with lowest penalty.

\[
\text{Penalty}(\text{REG}_i) = \alpha \times \text{proifling\_penalty}(\text{REG}_i) \\
+ (1- \alpha) \times \text{history\_penalty}(\text{REG}_i) \\
0 < \alpha < 1
\]
Outline

• Introduction
• Motivation
• The Proposed Methods
  – Binding Method
  – Forwarding Method
• Experimental Results
• Conclusions
Forwarding Method – Motivation

- The forwarding unit *exists* in a pipeline architecture.

- Useless data is *still read* from register file when forwarding occurs.
Forwarding Method
– Main Idea

• Forwarding condition is detected at compile time and encoded into instructions.

• Sub-bank register file is determined to be turned off or not at run time.

• A Forwarding-aware Scheduling Algorithm is proposed to allow more operand-forwarding.
### The Modified Bundle Format

#### Flag encoding

<table>
<thead>
<tr>
<th>Flag encoding</th>
<th>Forwarding condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>no forwarding occurs</td>
</tr>
<tr>
<td>001</td>
<td>Inst₀, 1ˢᵗ source register</td>
</tr>
<tr>
<td>010</td>
<td>Inst₁, 1ˢᵗ source register</td>
</tr>
<tr>
<td>011</td>
<td>Inst₂, 1ˢᵗ source register</td>
</tr>
<tr>
<td>100</td>
<td>Conflict</td>
</tr>
<tr>
<td>101</td>
<td>Inst₀, 2ⁿᵈ source register</td>
</tr>
<tr>
<td>110</td>
<td>Inst₁, 2ⁿᵈ source register</td>
</tr>
<tr>
<td>111</td>
<td>Inst₂, 2ⁿᵈ source register</td>
</tr>
</tbody>
</table>

The format consists of 41 bits per instance and 5 temporary bits. The flag indicates the forwarding condition.
The Modified Datapath

- EXP
  - Flag [2:0] Decoder
  - Inst. Temp Decoder

- REN
  - Inst. Decoder

- REG
  - Reg File
  - Reg File
  - Reg File

- EXE
  - Forwarding Control Unit
  - MUX
  - MUX

Diagram showing the modified datapath with various components and connections between them.
Outline

• Introduction
• Motivation
• The Proposed Methods
  – Binding Method
  – Forwarding Method
• Experimental Results
• Conclusions
Experiment Flow

- **PowerStone**
  - Benchmark
  - Compiler Simulator
  - Thermal Simulator
  - Temperatur e Result

- **PowerImpact**
  - Power Configuration
  - Compiler Simulator

- **Trimaran**
  - Floorplan

- **Itaninum Processor**
  - HotSpot 4.0
Outline

- Introduction
- Motivation
- The Proposed Methods
  - Binding Method
  - Forwarding Method
- Experimental Results
- Conclusions
Conclusions

• Binding algorithm and forwarding method are proposed to reduce the hotspot of register file.

• The peak temperature reduction reach 7.89°C in the best case and 7.22°C in average.
THANK YOU !!
Profiling_penalty

- To represent the thermal relationship between register file and its adjacent units.

\[
\text{profiling\_penalty}(\text{REG}_1) = (U_{1\text{Temp}} \times w_1) + (U_{2\text{Temp}} \times w_2) + (U_{4\text{Temp}} \times l_1) + (U_{5\text{Temp}} \times w_5)
\]
History_penalty

• To represent the thermal relationship within sub-bank register files.
• The CFG (Control Flow Graph) is used to accumulate the access counts of register file.

Basic block $\alpha$
$C(\alpha) = 10$
$\Rightarrow \alpha$ executes 10 cycles
$\Rightarrow$ access counts of $\text{REG}_1 = 2$
$\Rightarrow$ access counts of $\text{REG}_2 = 3$
History_penalty

• To represent the thermal relationship within sub-bank register files.

\[
\begin{align*}
C(\alpha) & = 10 \\
C(\beta) & = 30 \\
C(\beta') & = 20 \\
\text{REG}_1 & = 2 \\
\text{REG}_2 & = 2 \\
\text{REG}_1 & = 5 \\
\text{REG}_2 & = 3 \\
\end{align*}
\]

\[
\begin{align*}
\text{Prob}(\text{edge } \alpha \beta) & = 0.4 \\
\text{Prob}(\text{edge } \alpha \beta') & = 0.6 \\
\end{align*}
\]

\[
\begin{align*}
priority(\beta) &= \text{Prob}(\text{edge } \alpha \beta) \times (1 + \text{REG}_1) \\
&= 0.6 \times (1+2) = 1.8 \\
priority(\beta') &= \text{Prob}(\text{edge } \alpha \beta') \times (1 + \text{REG}_1) \\
&= 0.4 \times (1+5) = 2.4 \\
\text{Count}(\text{REG}_1) &= 2 + 5 = 7 \\
\end{align*}
\]

\[C(\alpha) + C(\beta') > \text{threshold cycle?}\]
## Performance Penalty

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Orig. Cycle count</th>
<th>Binding Pen (%)</th>
<th>Forwarding Pen (%)</th>
<th>Combined Pen (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bilv</td>
<td>376064357</td>
<td>0.0</td>
<td>1.3</td>
<td>0.9</td>
</tr>
<tr>
<td>blat</td>
<td>50440416</td>
<td>0.0</td>
<td>1.7</td>
<td>1</td>
</tr>
<tr>
<td>crc</td>
<td>5660433</td>
<td>0.0</td>
<td>0.7</td>
<td>0.5</td>
</tr>
<tr>
<td>des</td>
<td>41964763</td>
<td>0.0</td>
<td>1.5</td>
<td>0.8</td>
</tr>
<tr>
<td>energ</td>
<td>711784</td>
<td>0.0</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>fir</td>
<td>45077873</td>
<td>0.0</td>
<td>1.6</td>
<td>1.1</td>
</tr>
<tr>
<td>summin</td>
<td>492696</td>
<td>0.0</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>whet</td>
<td>77202731</td>
<td>0.0</td>
<td>1.5</td>
<td>0.9</td>
</tr>
<tr>
<td><strong>average</strong></td>
<td><strong>0.0</strong></td>
<td><strong>1.3</strong></td>
<td><strong>0.9</strong></td>
<td></td>
</tr>
</tbody>
</table>