

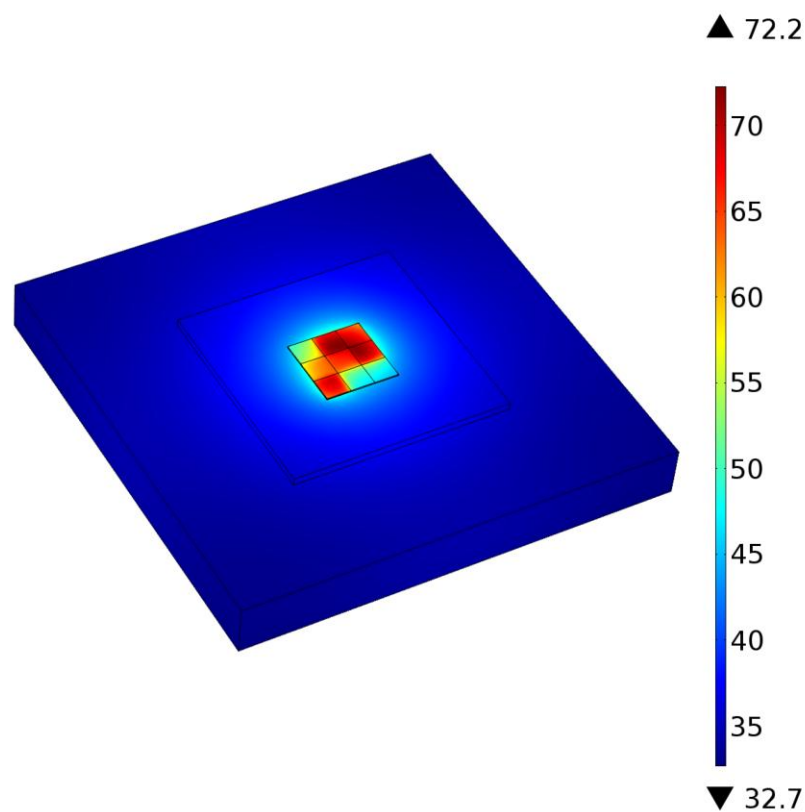


# 高性能众核芯片动态热管 理技术研究

马健

# 为什么需要动态温度管理

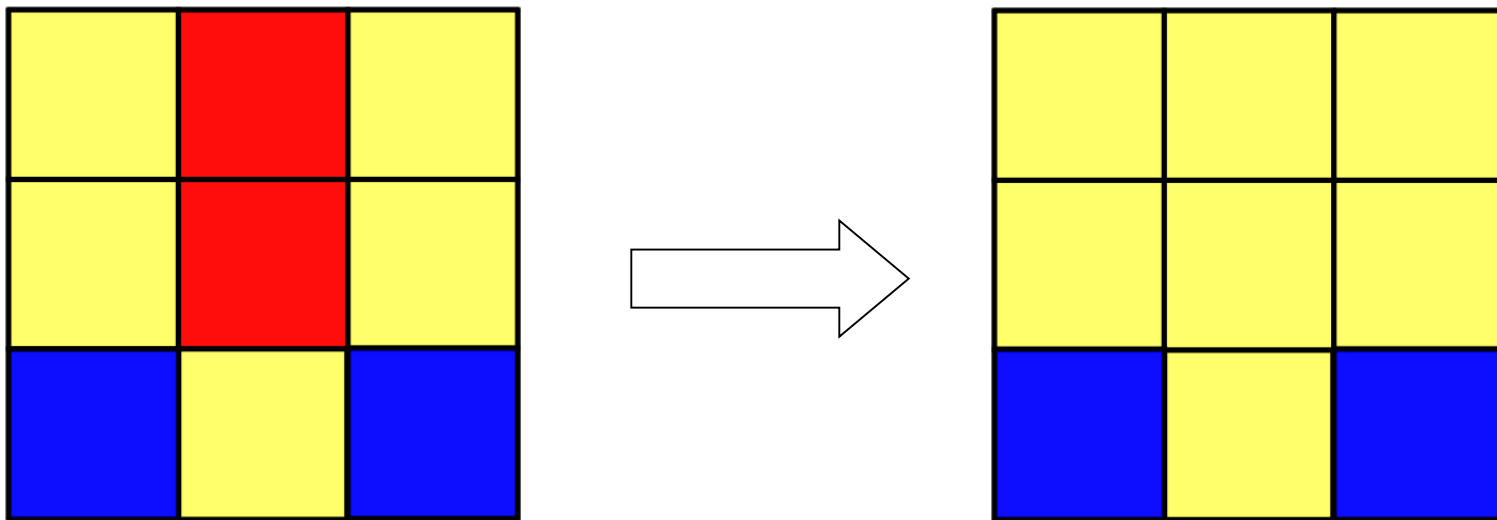
- 功耗密度不断增长导致高温热问题，在众核领域，高温点问题更加严重
  - 芯片高温或局部高温影响系统可靠性
  - 高温使静态功耗增加
  - 热问题使系统冷却开销很大



多核芯片的温度分布

# 动态温度管理技术

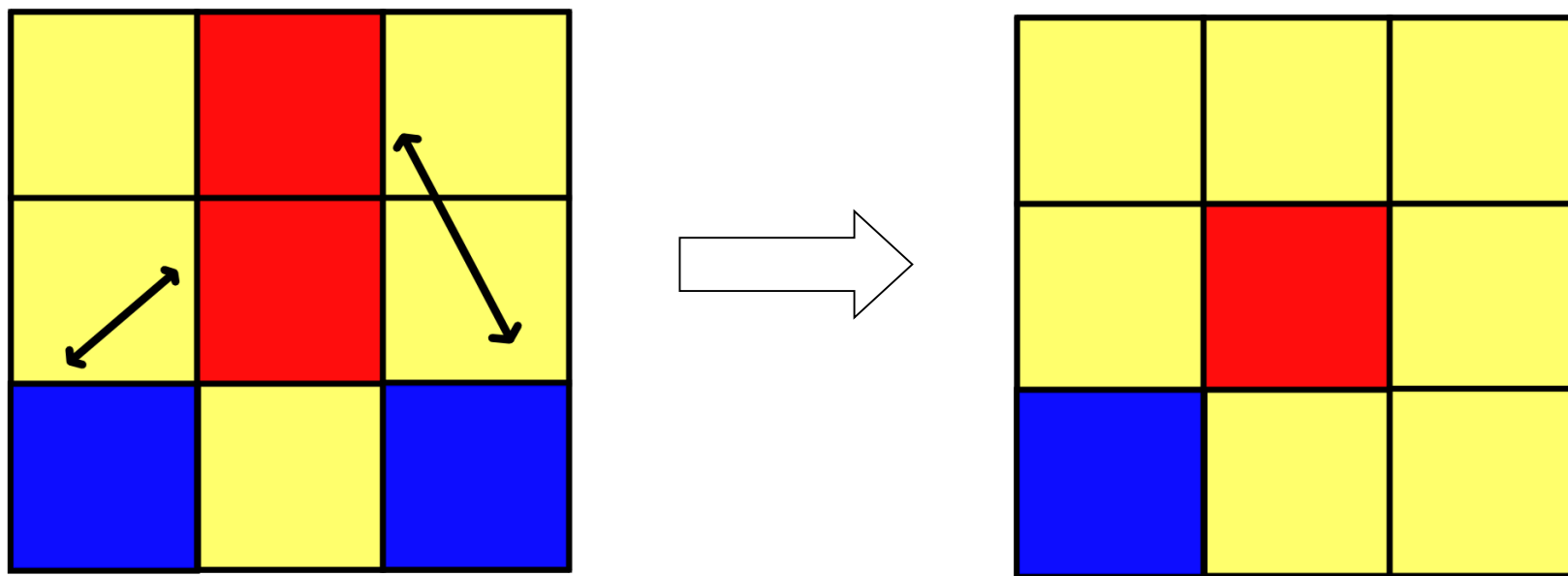
- 动态电压频率调整技术 (DVFS):
  - $P \sim f v^2$
  - 通过调整电压和频率来降低动态功耗
  - 温度降低的同时处理速度下降



红色核： 温度超出安全温度  
黄色核： 温度刚刚在安全温度以下  
蓝色核： 温度远低于安全温度

# 动态温度管理技术

- 任务迁移：
  - 将高温核上的重负载任务迁出，来避免高温问题



# 热模型

$$T(k+1) = AT(k) + B_d P(k)$$

$$Y(k) = LT(k)$$

$T$  表示温度向量，包括处理器核的温度和其他部分。

$Y$  表示输出向量，即核的温度向量。

$P$  表示输入向量，即核的功耗向量

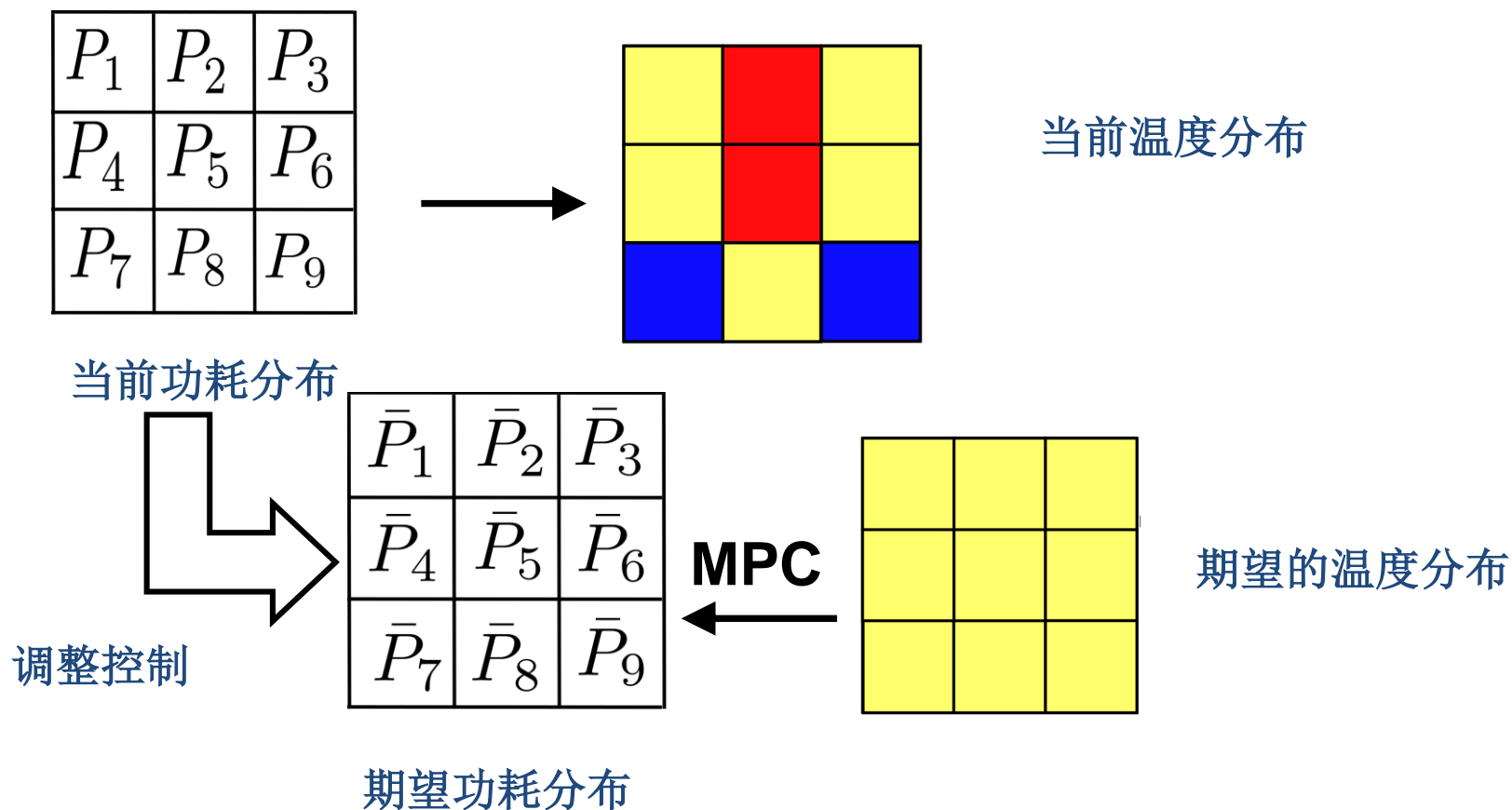
$A$   $B_d$  表示hotspot热模型提取的热容热阻信息的离散化形式

$L$  是一个选择矩阵，即从  $T$  中将核的温度选择出来

# 最近的动态温度管理方法

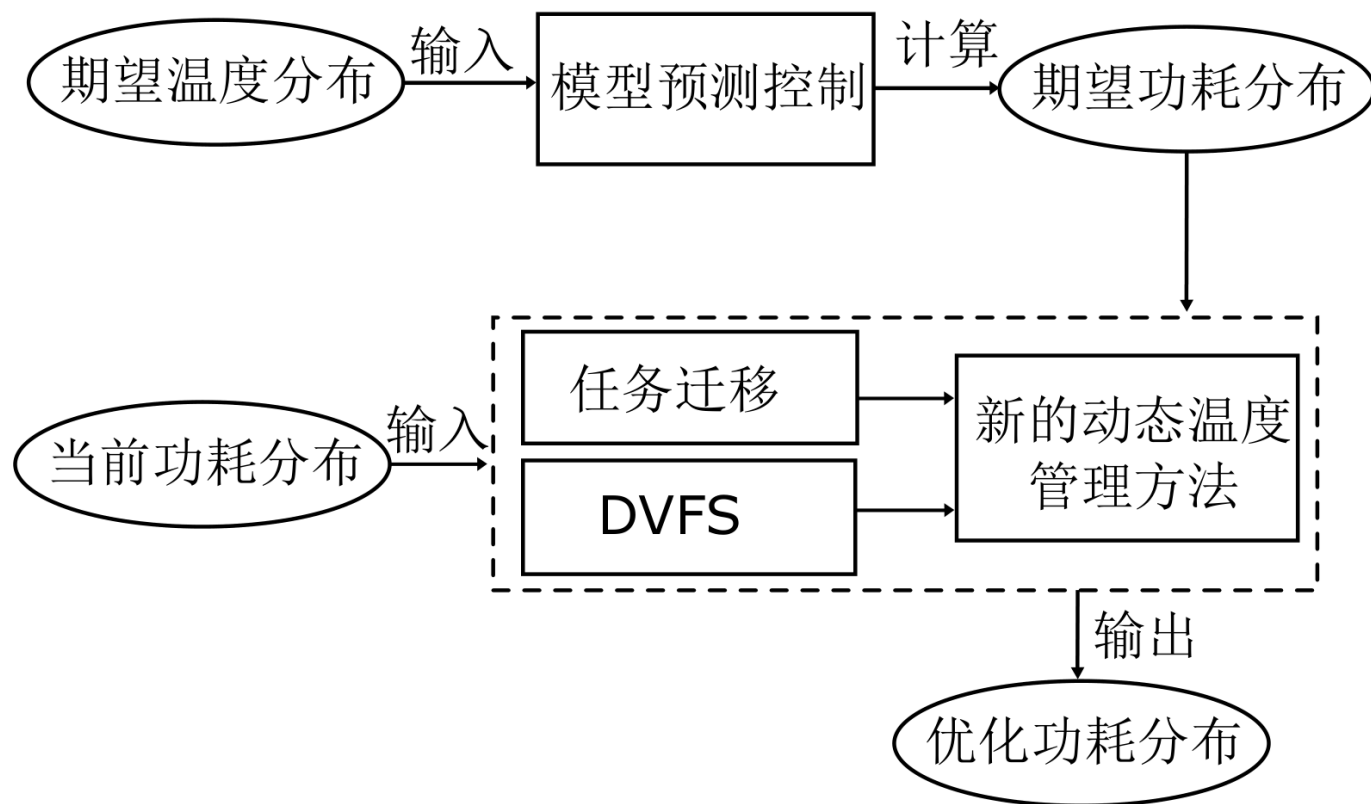
- 模型预测控制 (MPC)

- 提供指导性调整意见：计算出期望的功耗分布



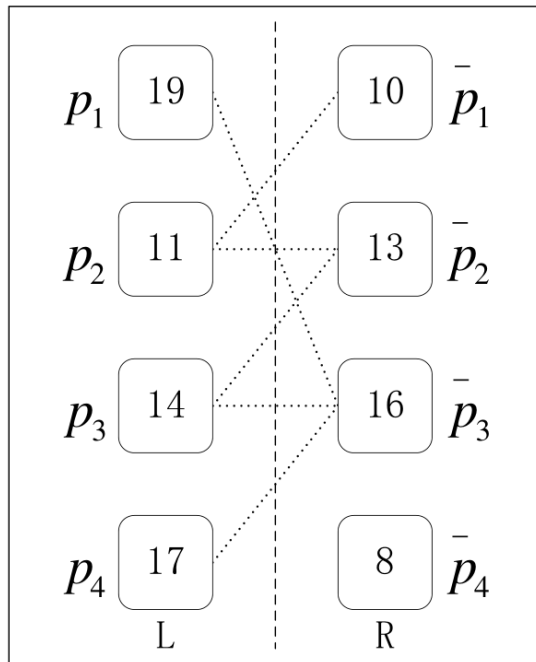
# 提出的混合方法

- 新的动态温度管理方法流程



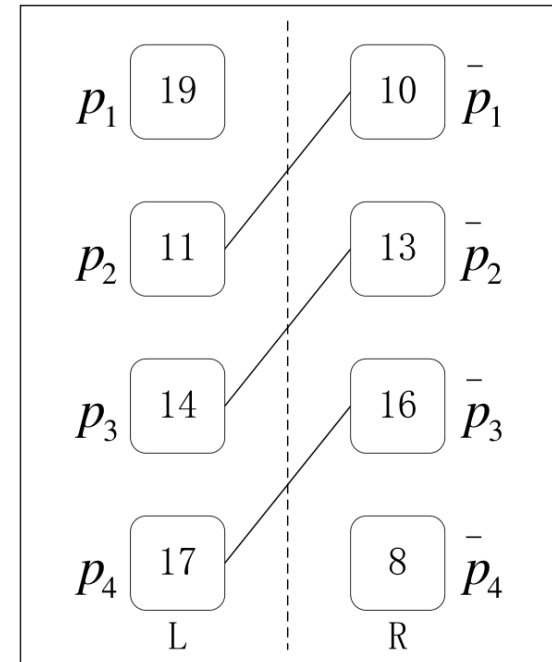
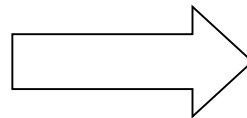
# 提出的混合方法

- 模型预测控制方法结合任务迁移和DVFS
  - 匈牙利算法：分配任务到正确的核（一定阈值下的二部图匹配）
  - DVFS保证匹配不上的核在安全温度以下



匹配前设定匹配阈值为3

利用匈牙利算法匹配



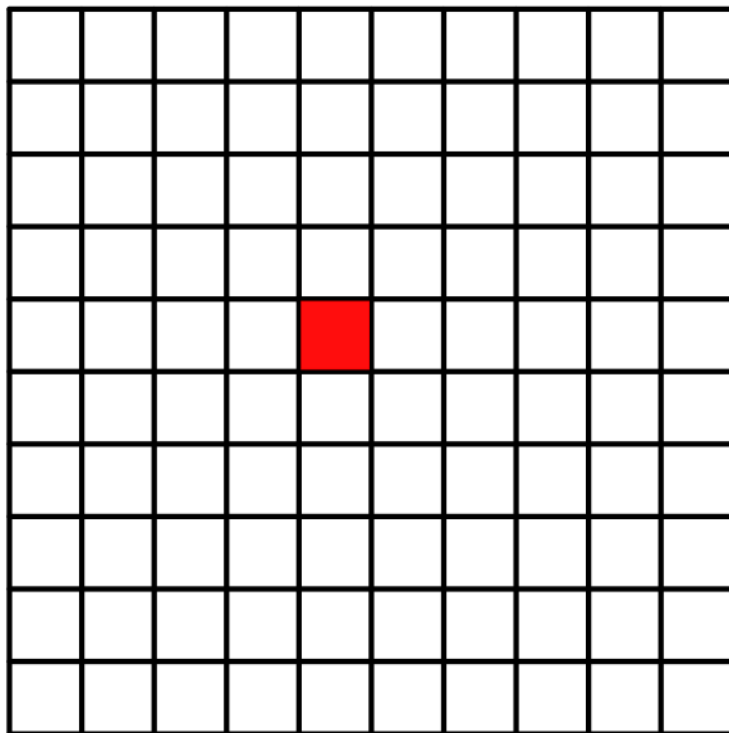
匹配后

4核匹配示例



# 提出的混合方法

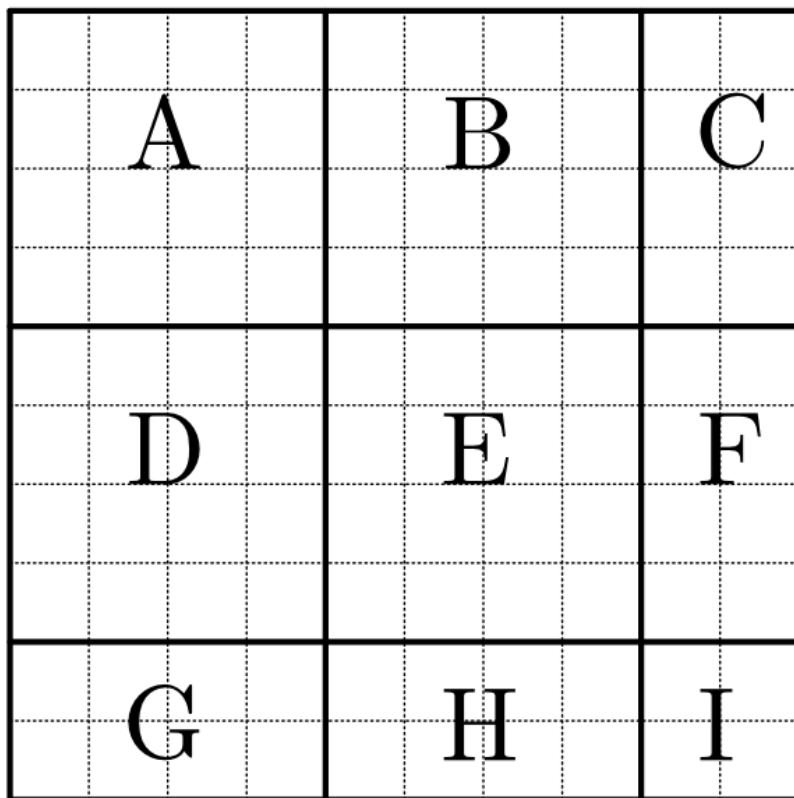
- 算法运算时间的问题
  - 随着核数增长，任务迁移决策的计算需要太多时间



100核芯片的例子

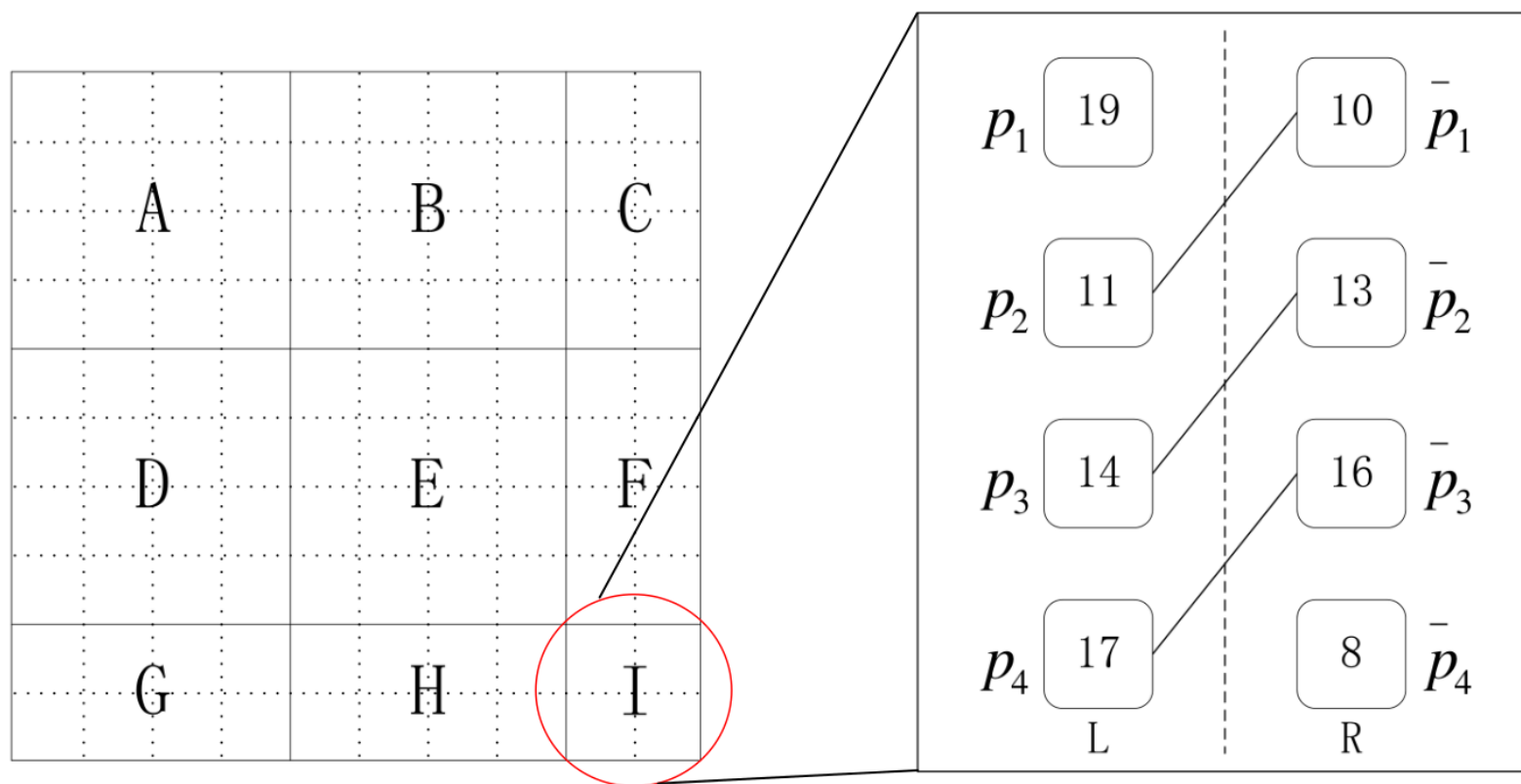
# 改进的分层方法

- 将芯片分块做两层任务迁移
  - 块内任务迁移
  - 块间任务迁移



# 改进的分层方法

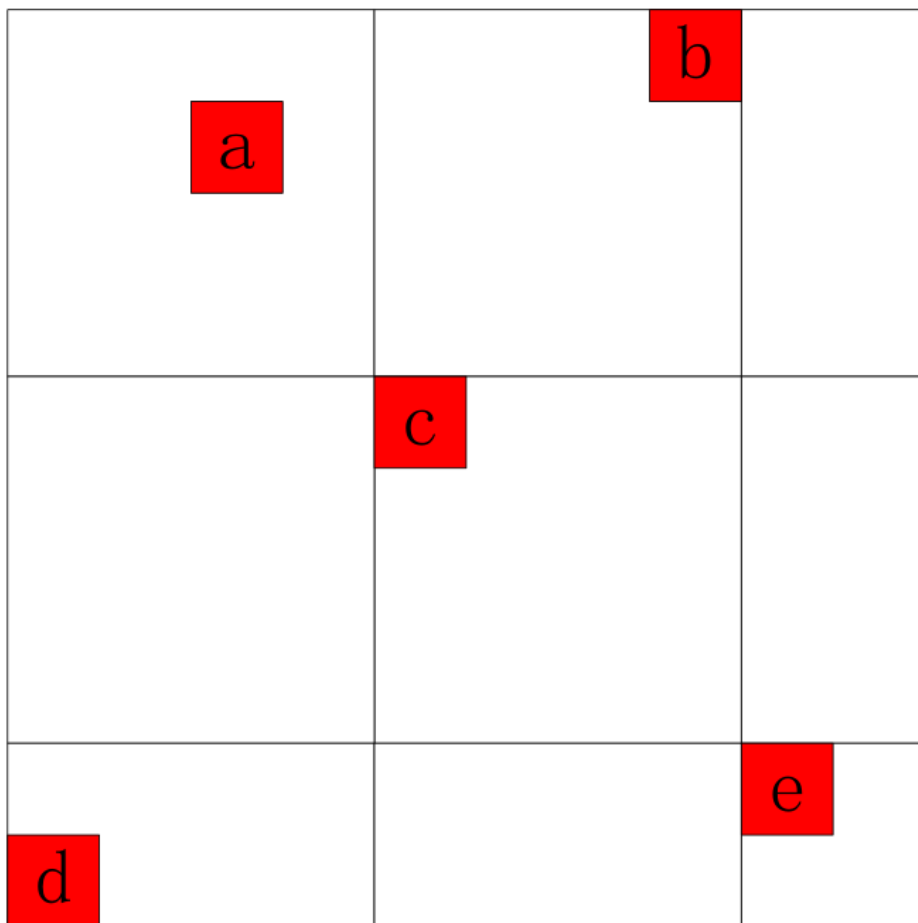
- 块内任务迁移



块I的例子

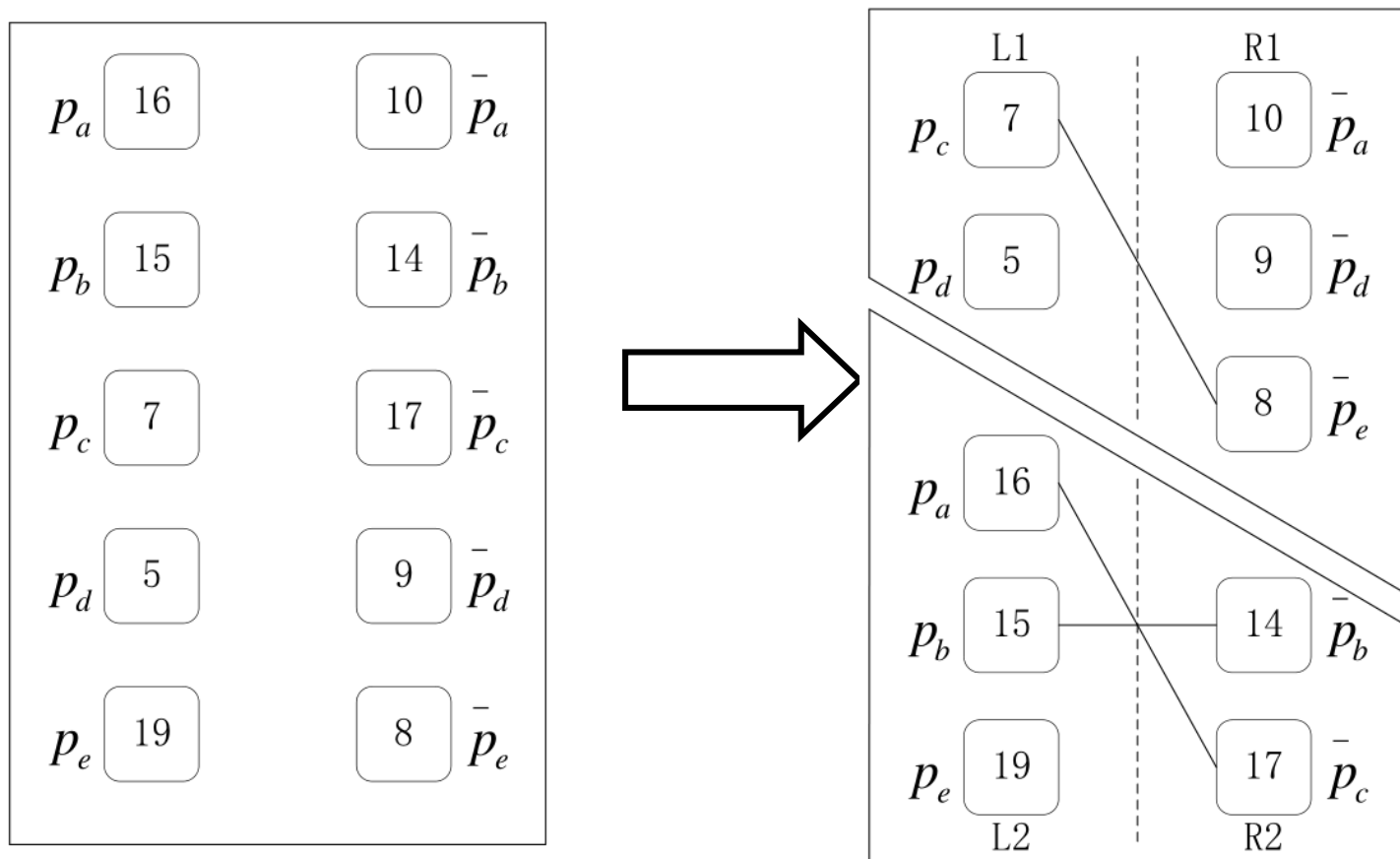
# 改进的分层方法

- 没有匹配上的核可能在其他块可以匹配
- 将未匹配核集中到一起



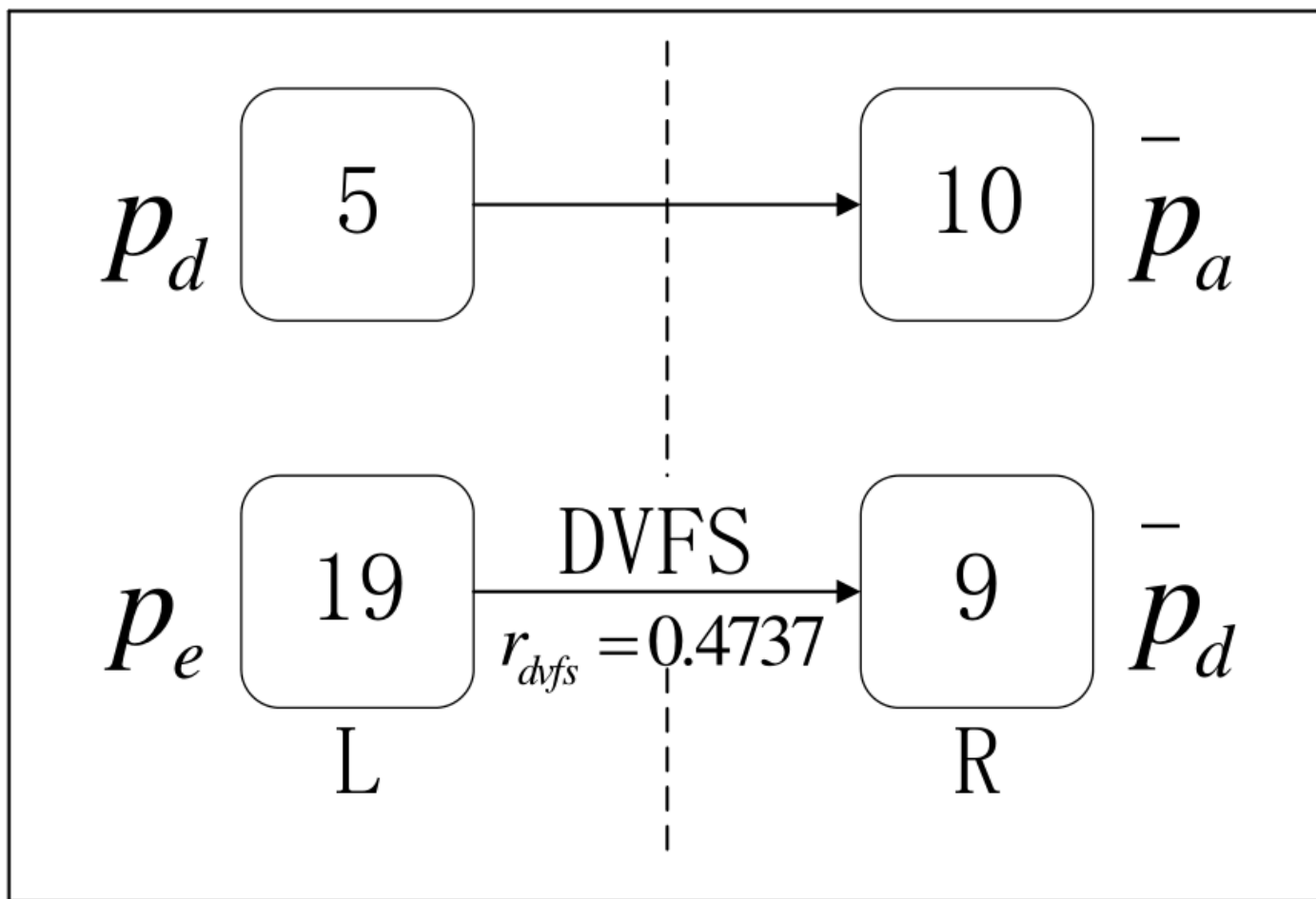
# 改进的分层方法

- 未匹配核的数量可能很大
  - 如果未匹配核数量很小，直接进行匹配.
  - 否则用最小割算法将其划分，每部分内部匹配



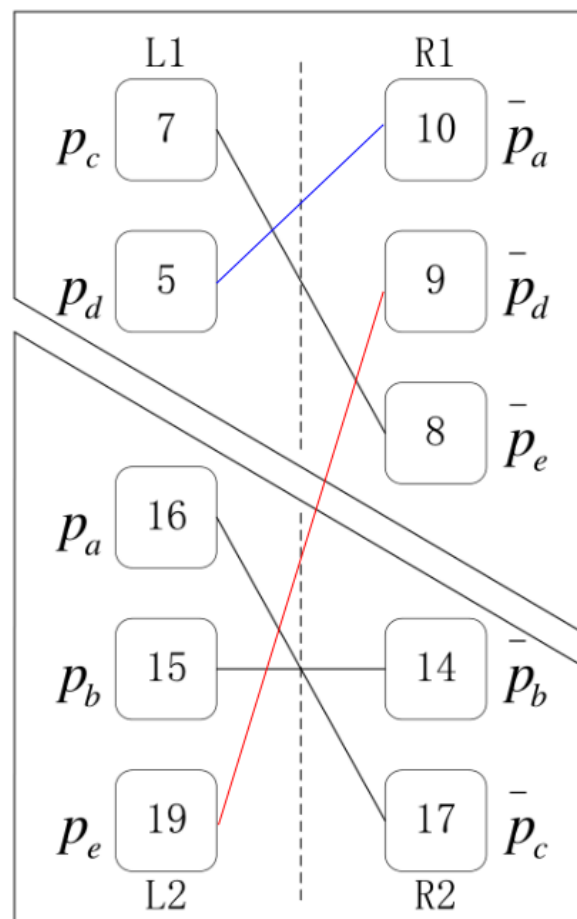
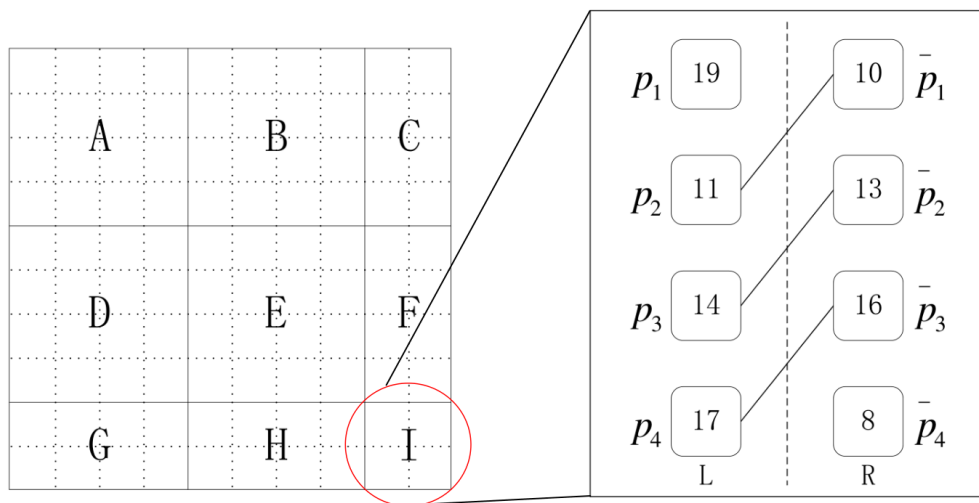
# 改进的分层方法

最后没有匹配上的核再采用DVFS处理。



# 改进的分层方法

- 功耗重新分配完成

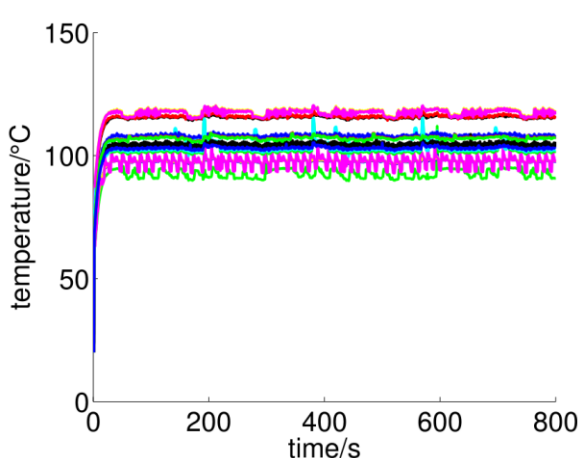


# 实验设定

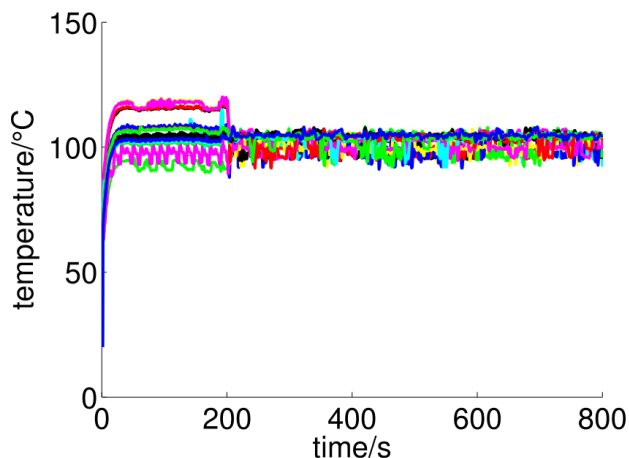
- 对于未分层的混合方法，设定16核，25核，36核与49核来进行测试比较。对于改进的分层方法设定了100核，256核，400核以及625核进行测试比较。
- 热模型用**HotSpot**得到。
- 功耗信息由**SPEC benchmarks** 在**Wattch**上运行得到。
- 环境温度设定为20摄氏度。
- 安全温度设定为105摄氏度。



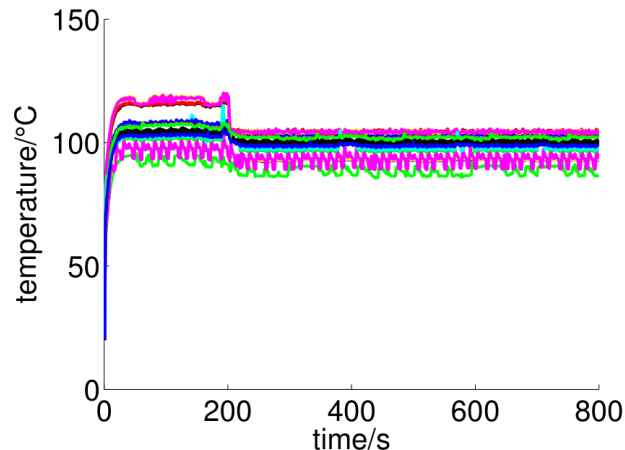
# 未分层混合方法的比较



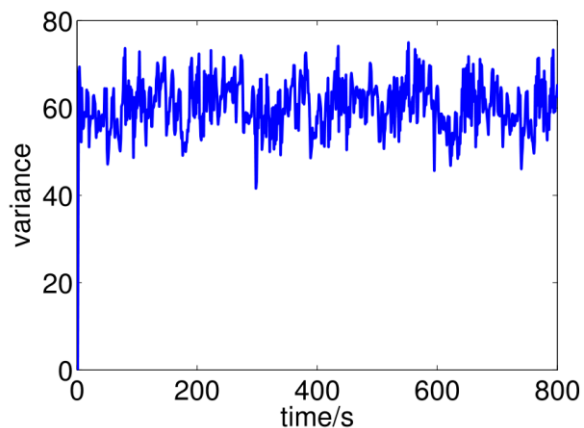
没有采用温度管理方法时的温度



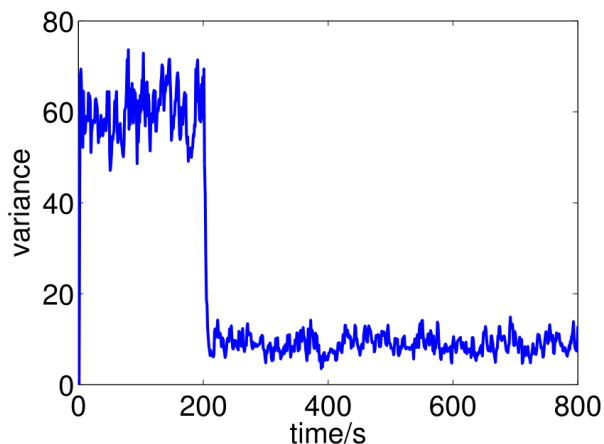
采用混合方法时的温度



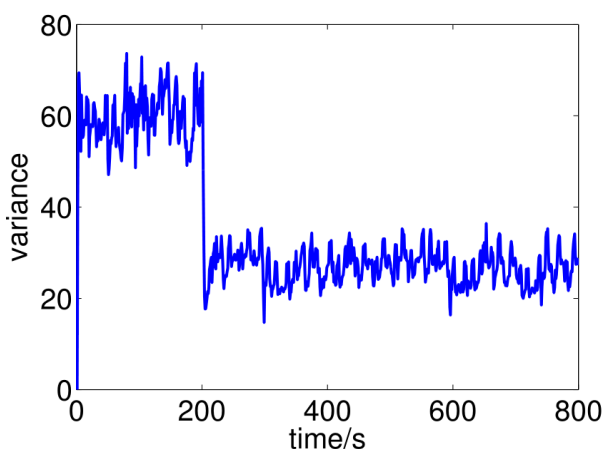
采用MPC只结合DVFS方法时的温度



没有采用温度管理方法时的核间温度方差



采用混合方法时的核间温度方差



采用MPC只结合DVFS方法时的核间温度方差

# 未分层混合方法的比较

核数	$MIPS_o$	周期 1s	
		$MIPS_d$	$MIPS_n$
16 核	651.5	630.6	643.8
25 核	530.2	508.0	520.4
36 核	477.1	460.4	471.4
49 核	442.6	426.4	439.1

$MIPS_o$ 表示理想状态下的每秒指令数

$MIPS_d$ 表示采用MPC只结合DVFS方法的每秒指令数

$MIPS_n$ 表示采用混合方法的每秒指令数

混合方法与其他方法的性能比较

# 未分层混合方法的比较

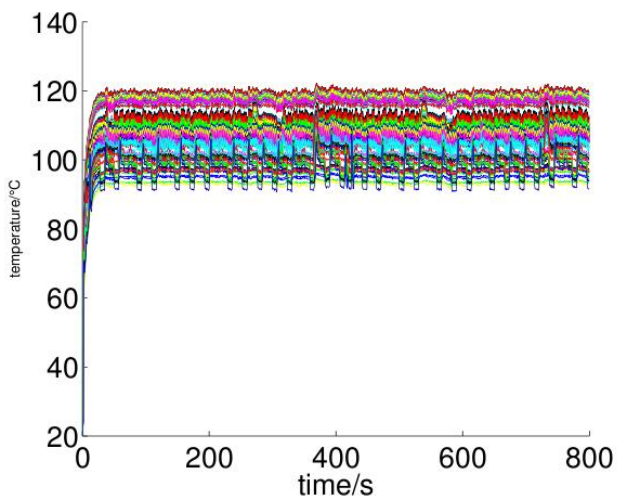
核数	周期1s	
	$t_p(10^{-3}s)$	$t_m(10^{-3}s)$
16 核	0.29	4
25 核	0.38	8
36 核	0.50	19
49 核	0.61	41

$t_p$  表示MPC部分的计算时间

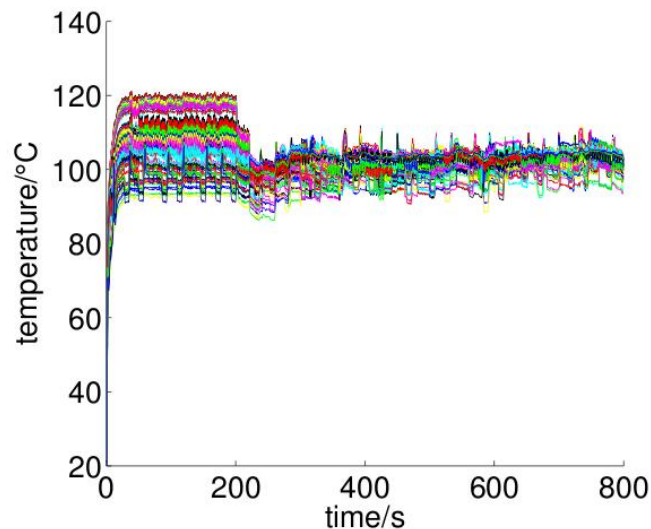
$t_m$  表示匹配部分的计算时间

混合方法与其他方法的计算时间比较

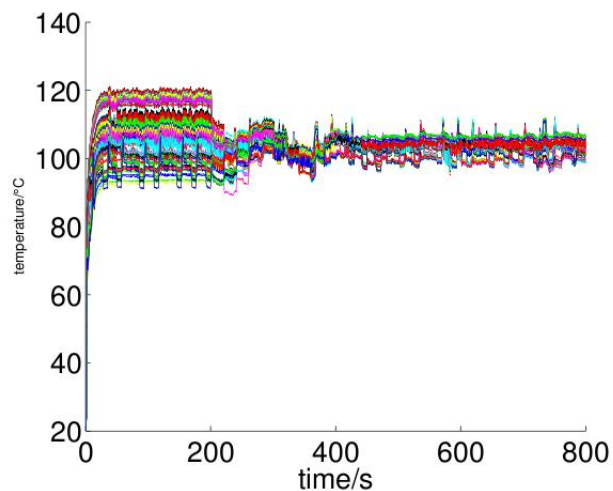
# 改进分层方法的比较



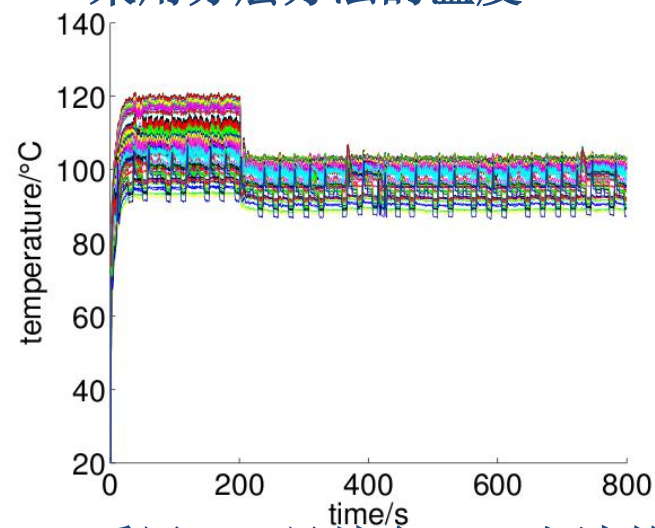
没有采用温度管理的温度



采用分层方法的温度

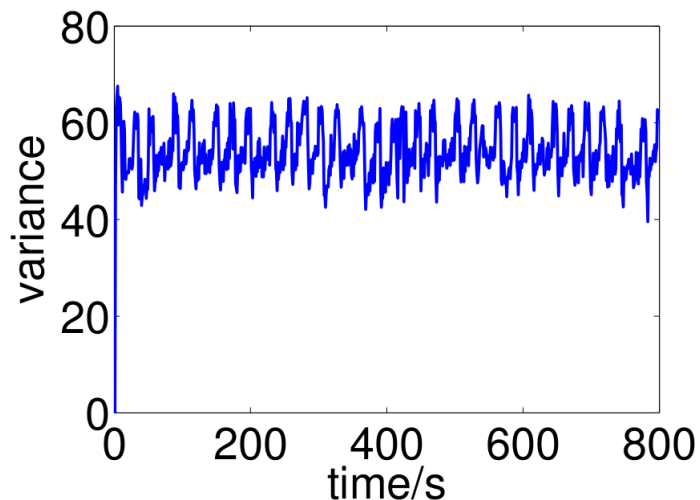


采用未分层方法的温度

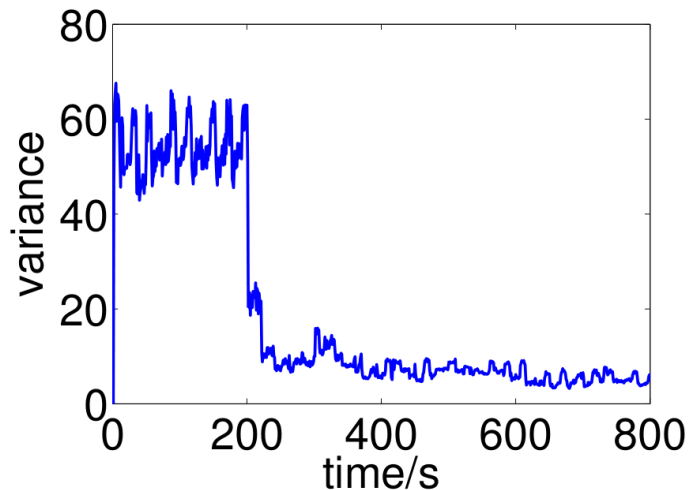


采用MPC只结合DVFS方法的温度

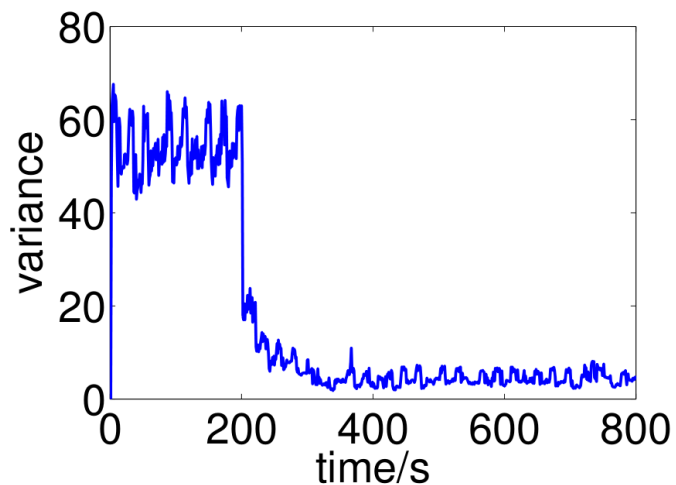
# 改进分层方法的比较



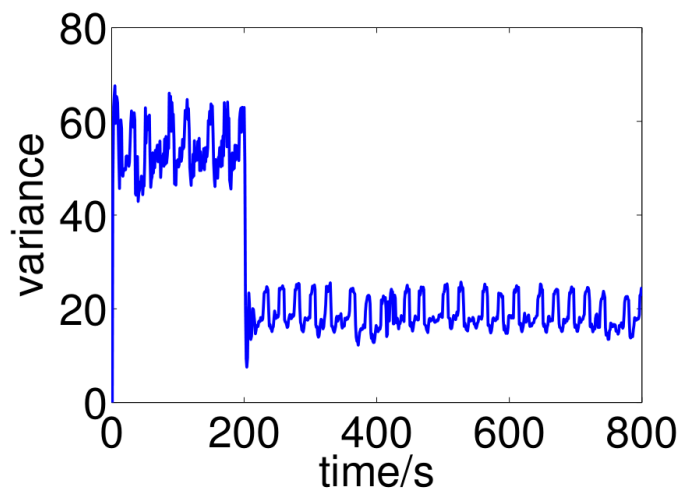
没有采用温度管理的核间温度方差



采用分层方法的核间温度方差



采用未分层方法的核间温度方差



采用MPC只结合DVFS方法的核间温度方差

# 改进分层方法的比较

	分层方法				未分层方法		
核数	$t_p$ ( $10^{-4}s$ )	$t_m$ ( $10^{-4}s$ )	$t_a$ ( $10^{-4}s$ )	$t_{all}$ ( $10^{-4}s$ )	$t_p$ ( $10^{-4}s$ )	$t_m$ (s)	$t_{all}$ (s)
100	1.58	1.63	0	3.21	1.49	0.01	0.01
256	2.85	19.26	1.58	23.7	2.93	0.45	0.45
400	4.88	7.77	3.27	15.9	5.38	1.90	1.90
625	9.09	12.27	17.49	38.8	8.27	8.63	8.63

分层方法与未分层方法的计算时间比较

$t_p$  表示MPC部分的计算时间

$t_m$  表示匹配部分的计算时间

$t_a$  表示最小割划分部分的计算时间

# 改进分层方法的比较

核数	$MIPS_o$	$MIPS_d$	$MIPS_h$
100 核 ( $10 \times 10$ )	290.8	279.6	281.5
256 核 ( $16 \times 16$ )	210.2	202.9	207.1
400 核 ( $20 \times 20$ )	182.1	174.7	178.8
625 核 ( $25 \times 25$ )	156.5	150.2	154.4

分层方法与其他方法的性能比较

$MIPS_o$  表示理想状态下的每秒指令数

$MIPS_d$  表示采用MPC只结合DVFS方法的每秒指令数

$MIPS_n$  表示采用混合方法的每秒指令数

# 总结

- 提出了模型预测控制结合任务迁移和**DVFS**的混合方法，并将其改进成分层方法
- 混合方法是首次将模型预测控制方法与任务迁移相结合，并将其化为任务分配问题解决。
- 针对任务分配决策时间长的问题，创新性地将任务分配问题分层处理，并引入最小割算法。
- 通过实验比较，新的方法具有更可靠的热控制，更大的可靠性和性能优势，改进分层之后减少了计算开销增强了扩展性。



谢谢!