Temperature Aware Floorplanning

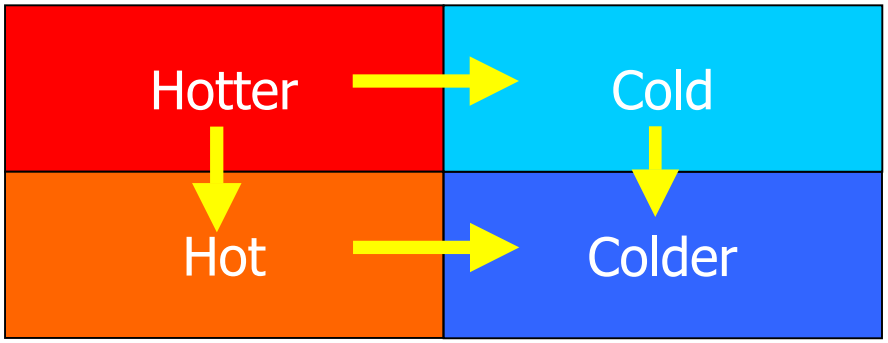
Yongkui Han, Israel Koren, C. Andras Moritz
ARTS Lab, ECE Dept
University of Massachusetts at Amherst

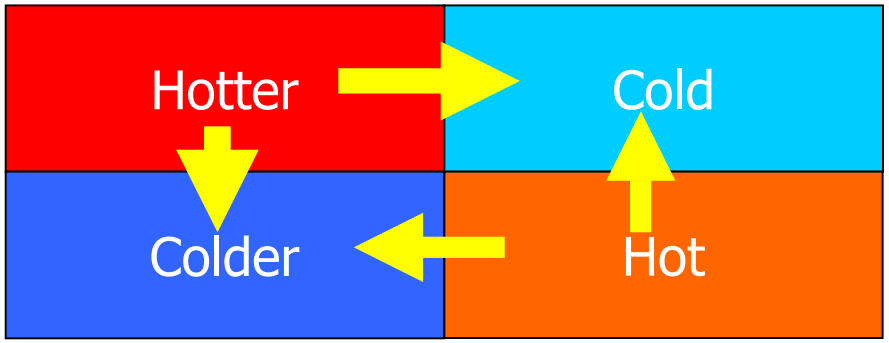
Temperature Aware Design

- Temperature can affect
  - **reliability**
  - **leakage power consumption**
  - **cost of cooling solutions**
- Temperature not uniform across blocks
- Temperature of a block depends on its power density and that of adjacent blocks
- Placement of blocks (floorplan) can affect maximum temperature of chip

Heat Diffusion Effects

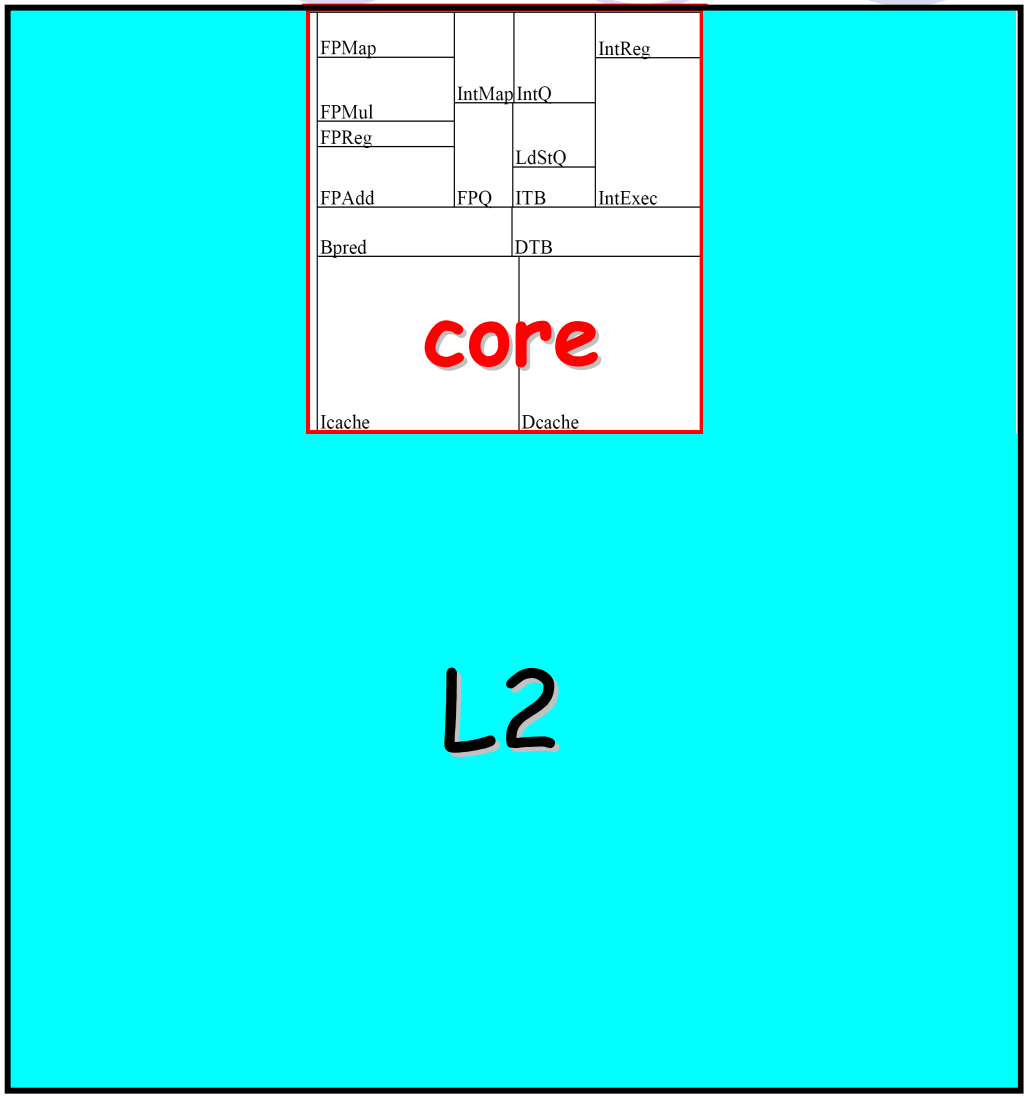
- Heat diffusion between adjacent blocks
- The larger the temperature difference, the larger the heat diffusion
- Beneficial to place hot blocks adjacent to cold blocks



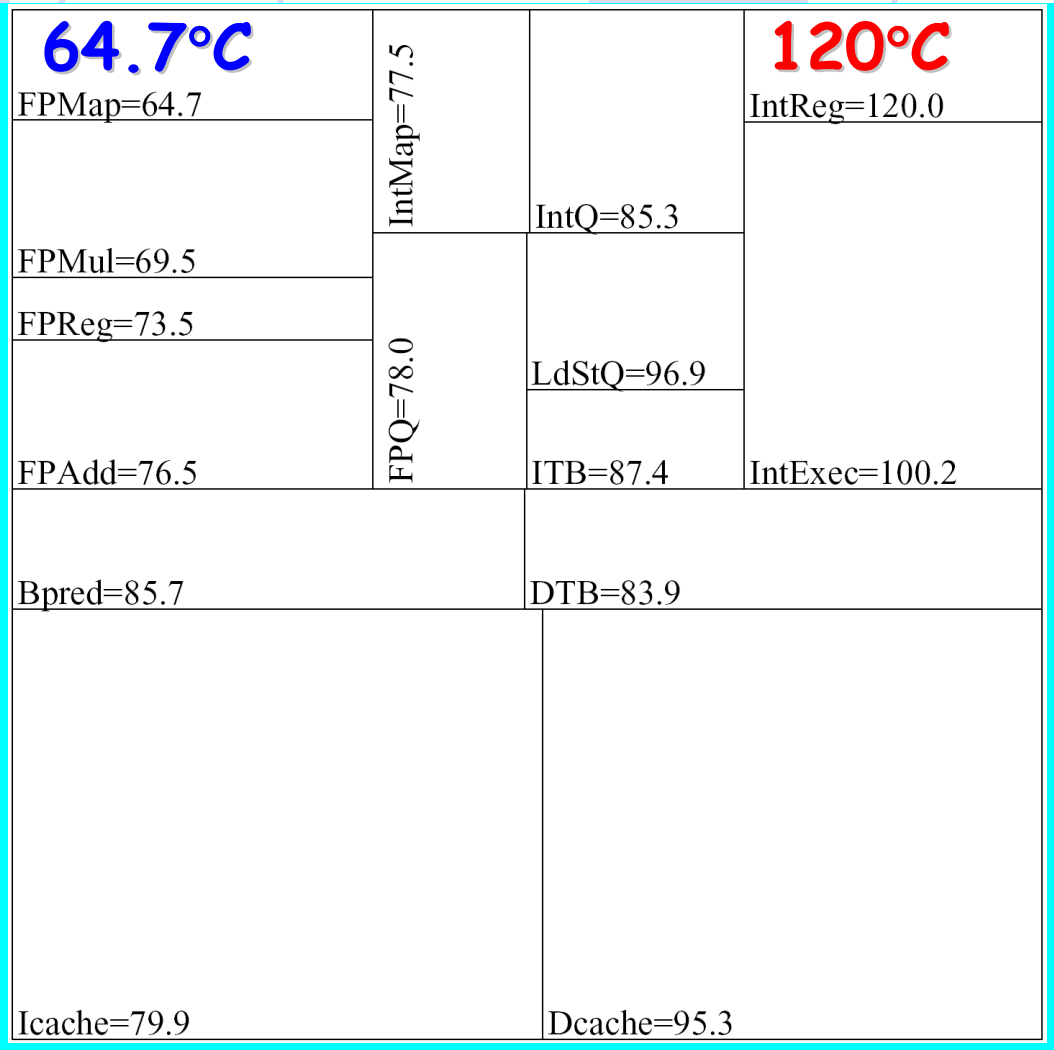


Temperature Aware Floorplanning Example

Original Alpha 21364 floorplan



Block Temperatures for *gcc* Benchmark

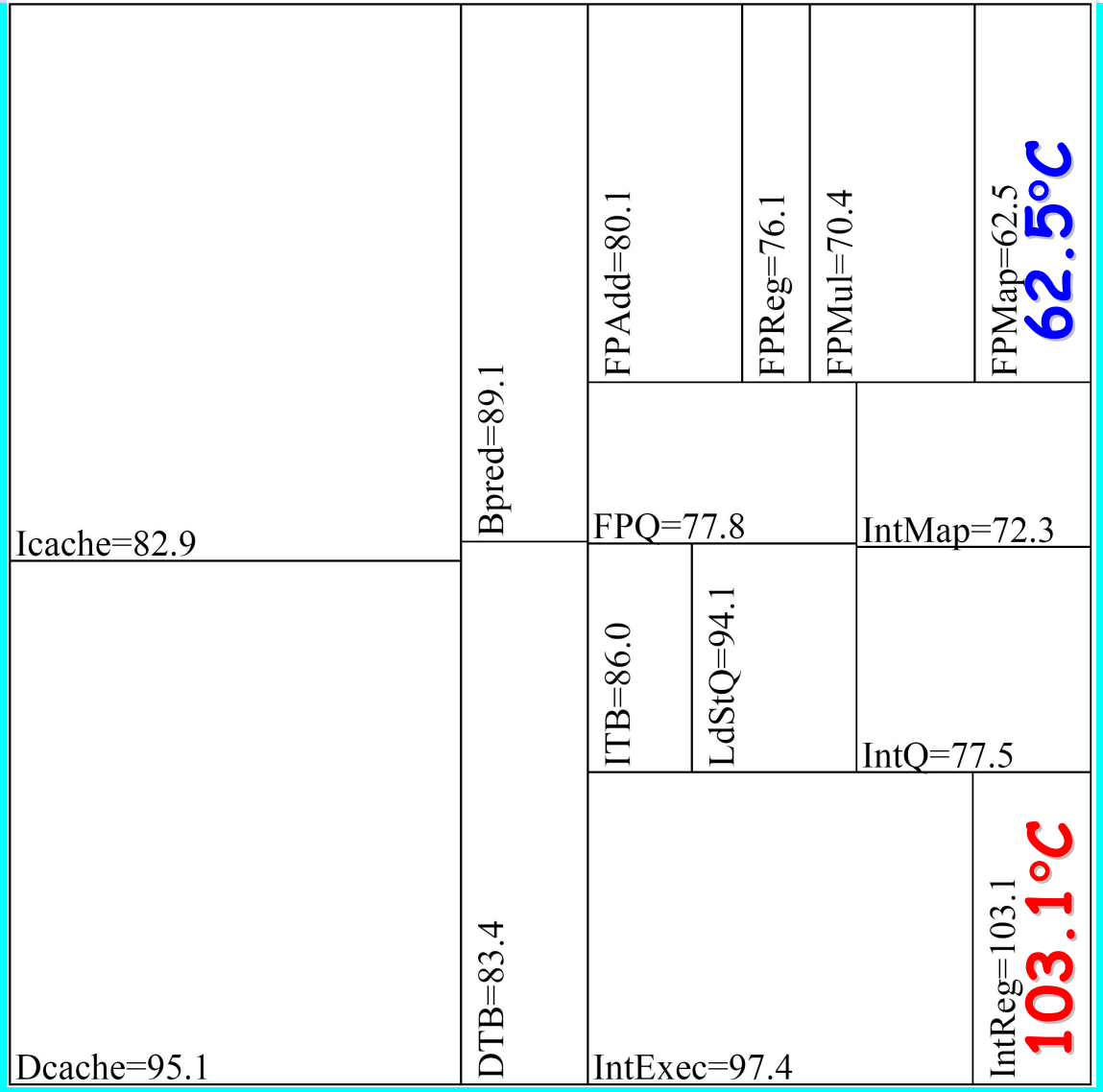


L2cache=43°C

Steady-state temperature calculated using HotSpot 2.0 (UVa)

Rotated Floorplan

L2cache=43°C

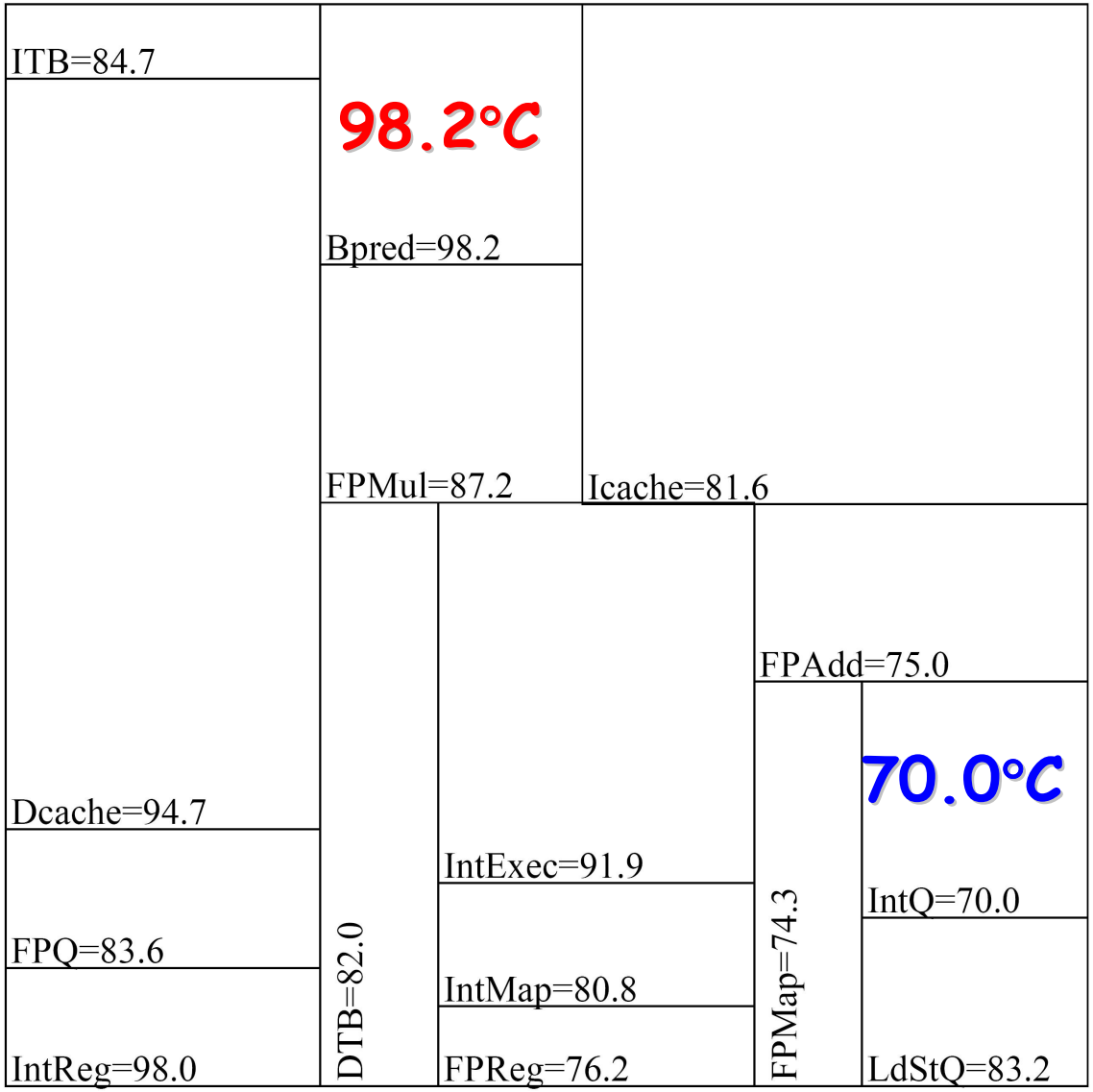


Reduction in MaxTemp: 16.9°C

Manually Generated Floorplan

Reduction in MaxTemp: 21.8°C

Spread of Temperature: 28.2°C vs 55.3°C

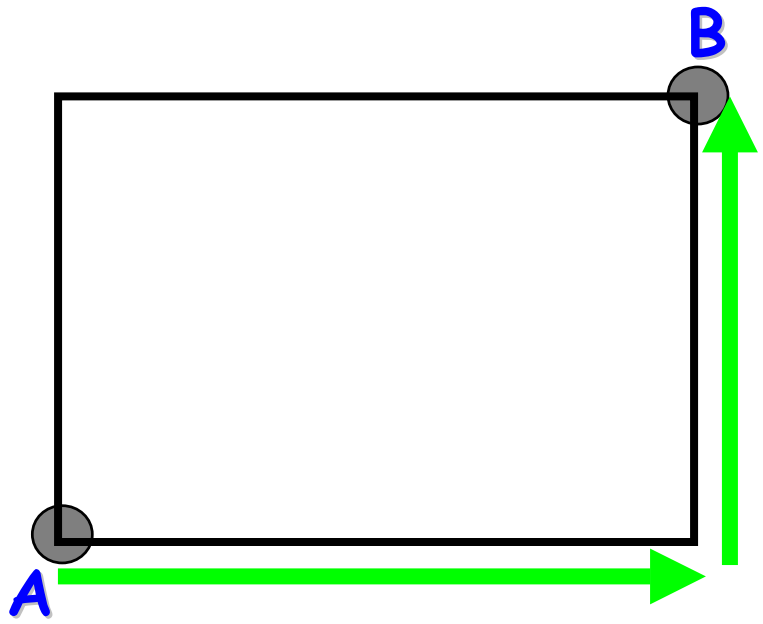


Chip Performance?

- These new floorplans can reduce the maximum temperature on chip
- Issues:
  - **How will these new floorplans affect chip performance?**
  - **If they compromise performance, can we find a floorplan which has both low maximum temperature and good performance?**
  - **What tradeoffs can we do in floorplanning?**

Metric of Chip Performance: Wire Length

- Wire length is a commonly used metric of chip performance
- HPWL (Half Perimeter Wire Length) is widely used in many floorplanning tools
- Critical paths are assigned larger weights



$$TotalWireLength = \Sigma W_{ij} * L_{ij}$$

Temperature Aware Floorplanning

- Parquet 2.0 floorplanning tool (UMich)
- Objective of Parquet:
  Minimize a linear combination of chip area and total wire length (using simulated annealing)

$$Ca * A + Cw * W$$

- Added temperature to objective function:

$$Ca * A + Cw * W + Ct * T$$

  - A - chip area, W - wire length, T - maximum temperature
  - We can adjust coefficients to achieve different optimization goals

- SPEC2000 benchmarks
  - Gcc as a representative benchmark

Approximation for the Temperature

- **Heat Diffusion**

$$H(d1, d2) = (d1 - d2) * shared\_length$$

  - **di – power density of block i**

$$H(d) = \Sigma H(d, di)$$

$$D = \Sigma H(d), \text{ for all } possibly\text{-}hot \text{ blocks}$$

Modified Floorplan: Low-temp

Low maximum temperature

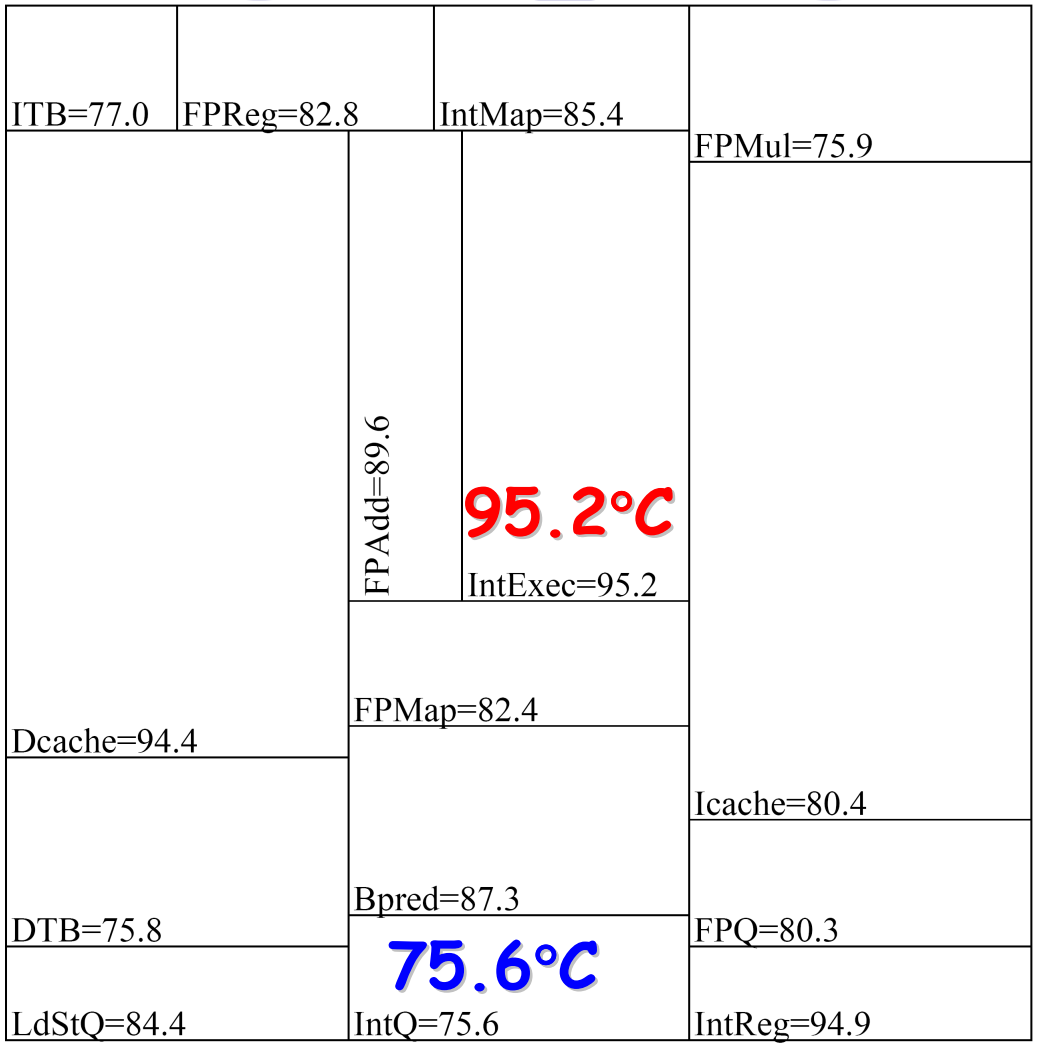
Reduction in MaxTemp: 24.8°C

Increase in Wirelength: 38.15%

Ca=.4, Cw=0, Ct=.6

Smaller spread of temperature: 19.6°C

Unacceptable wire length



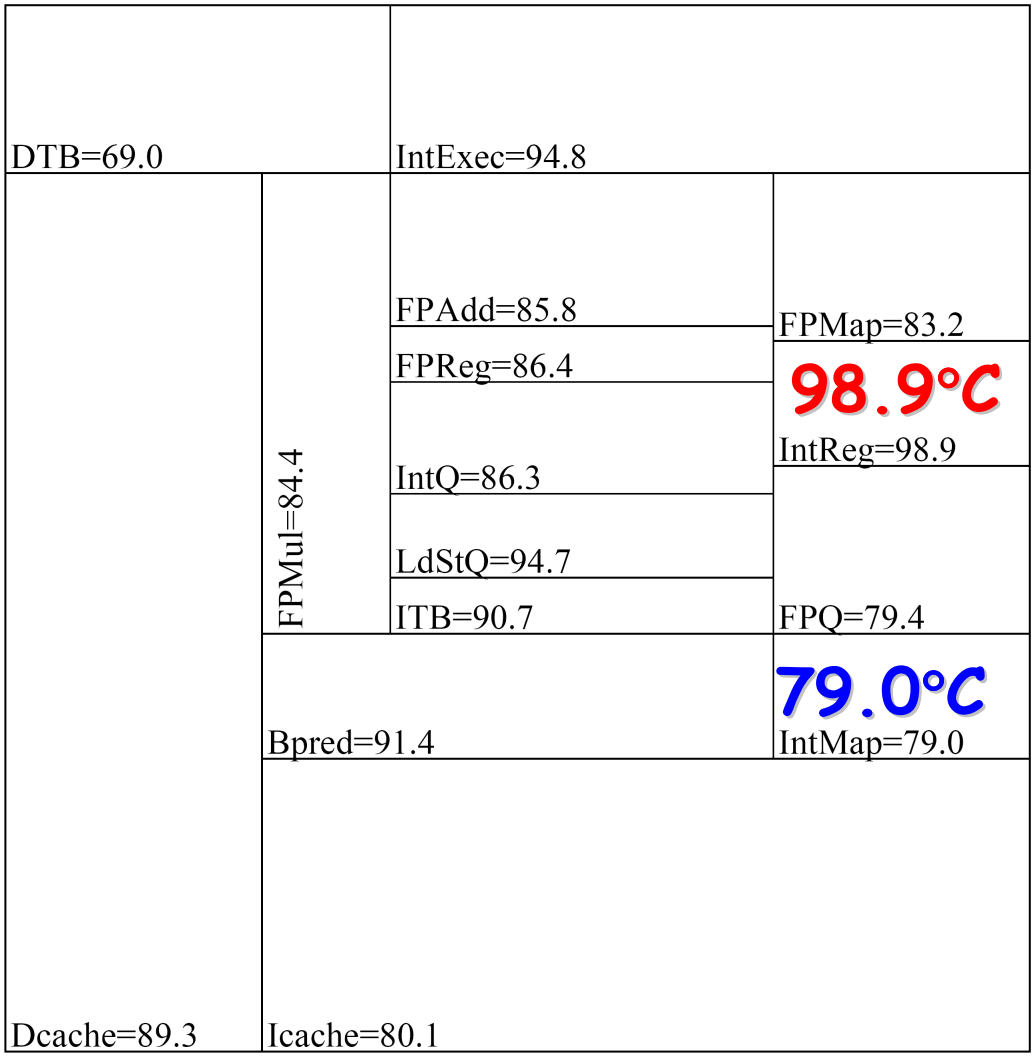
Modified Floorplan: Wire-temp

Short wirelength & low maximum temperature

Reduction in MaxTemp: 21.1°C
Increase in Wirelength: 0.67%

Ca=.3, Cw=.4, Ct=.3

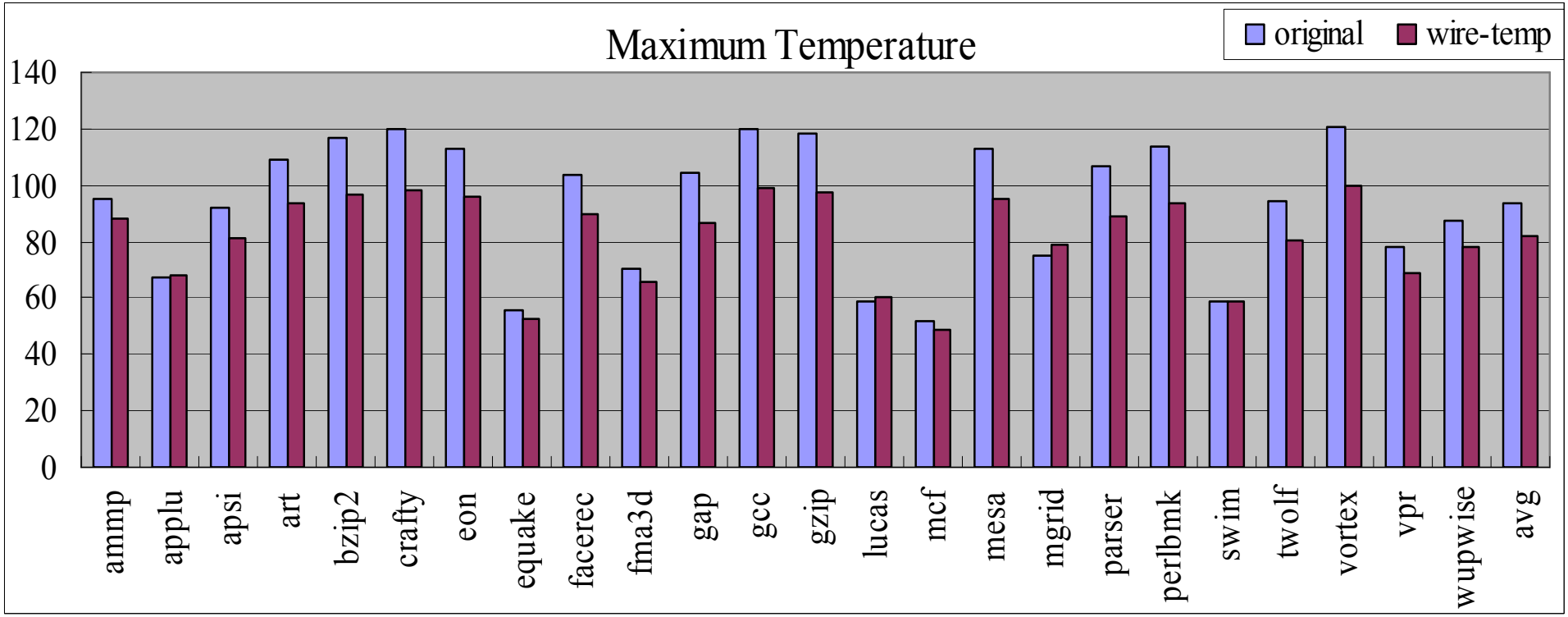
Small spread of temperature: 19.9°C



Temperature Reductions for SPEC2000 Benchmarks

Average MaxTemp reduction: 12°C

For many benchmarks, reductions are larger than 20°C



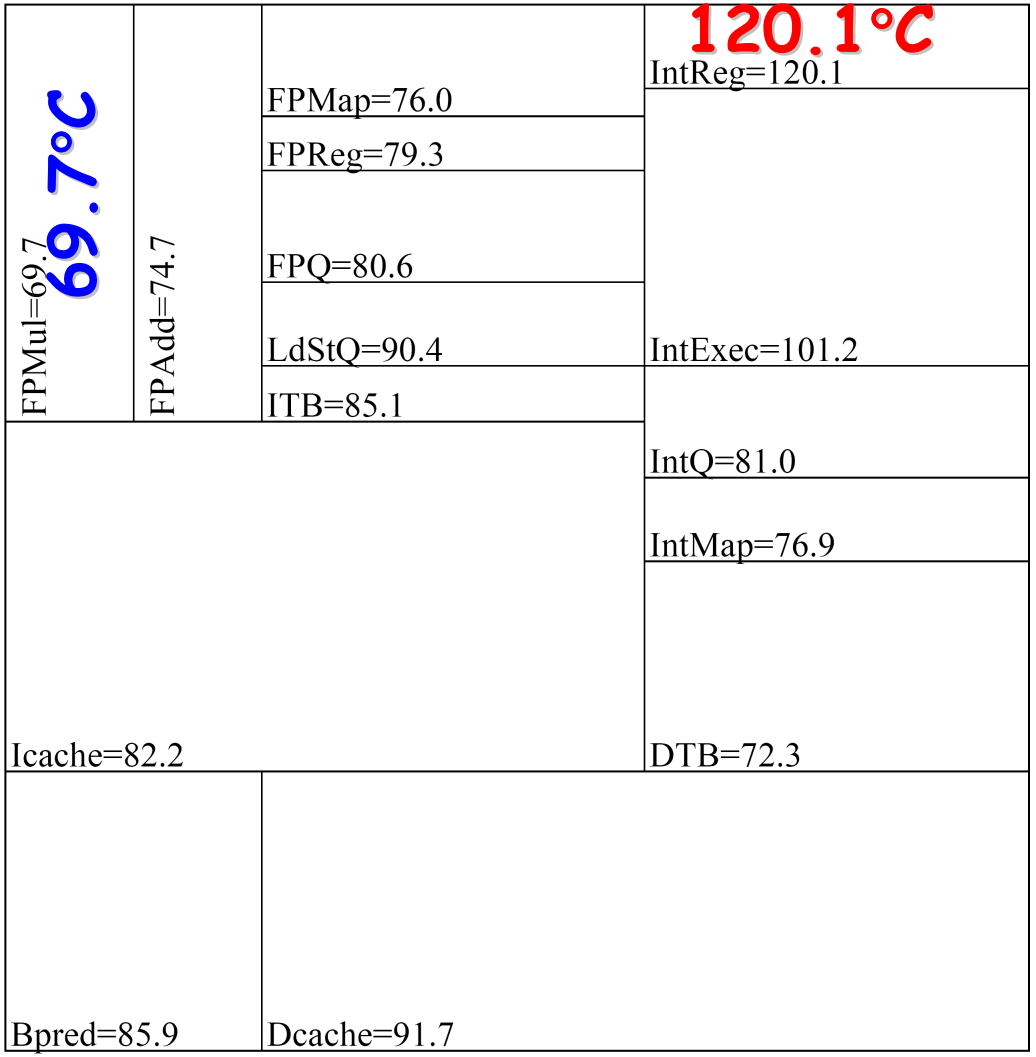
Modified Floorplan: Short-wire

Short wire length

Increase in MaxTemp: 0.1°C
Reduction in Wirelength: 4.07%

Ca=.4, Cw=.6, Ct=0

MaxTemp almost unchanged



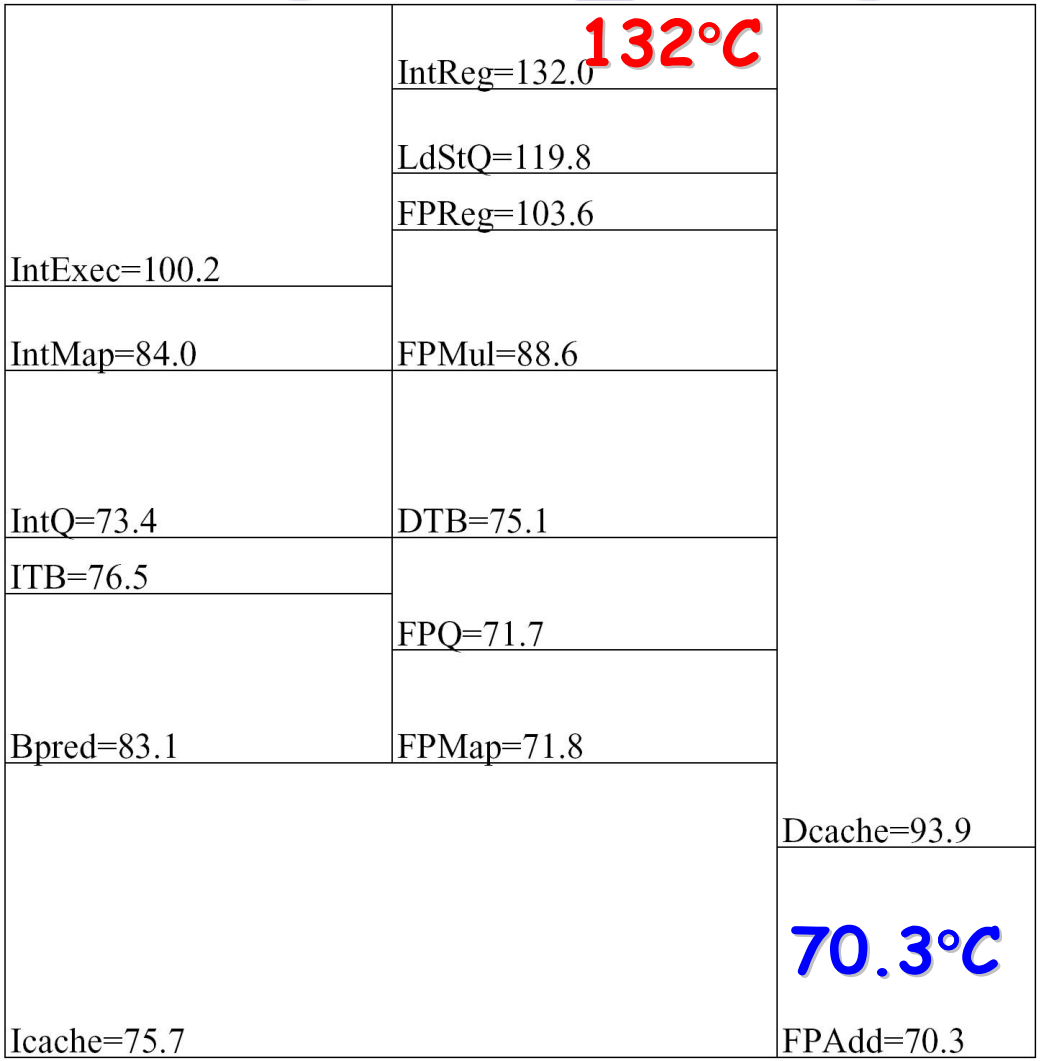
Modified Floorplan: High-temp

High maximum temperature

MaxTemp increase: 12.3°C
Wirelength increase: 6.36%

Ca=.4, Cw=0, Ct=-.6

High MaxTemp

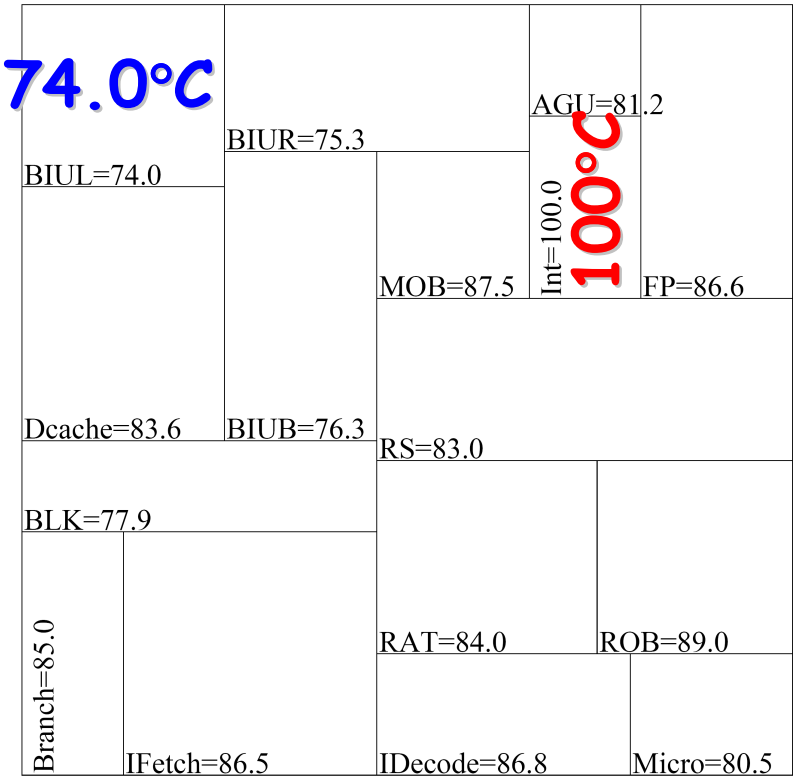


Summary - Alpha

Floorplan	Area (mm^2)	Increase	Wire length (m)	Increase	Temp ($°C$)	Reduction ($°C$)
Original	253.1	0%	17.93	0%	120.0	0
Manual	253.1	0%	23.21	29.45%	98.2	21.8
Rotated	253.1	0%	18.32	2.18%	103.1	16.1
Low-temp	254.1	0.4%	24.77	38.15%	95.2	24.8
Wire-temp	255.1	0.8%	18.05	0.67%	98.9	21.1
Short-wire	256.3	0.9%	17.20	-4.07%	120.1	-0.1
High-temp	255.1	0.8%	19.07	6.36%	132.3	-12.3

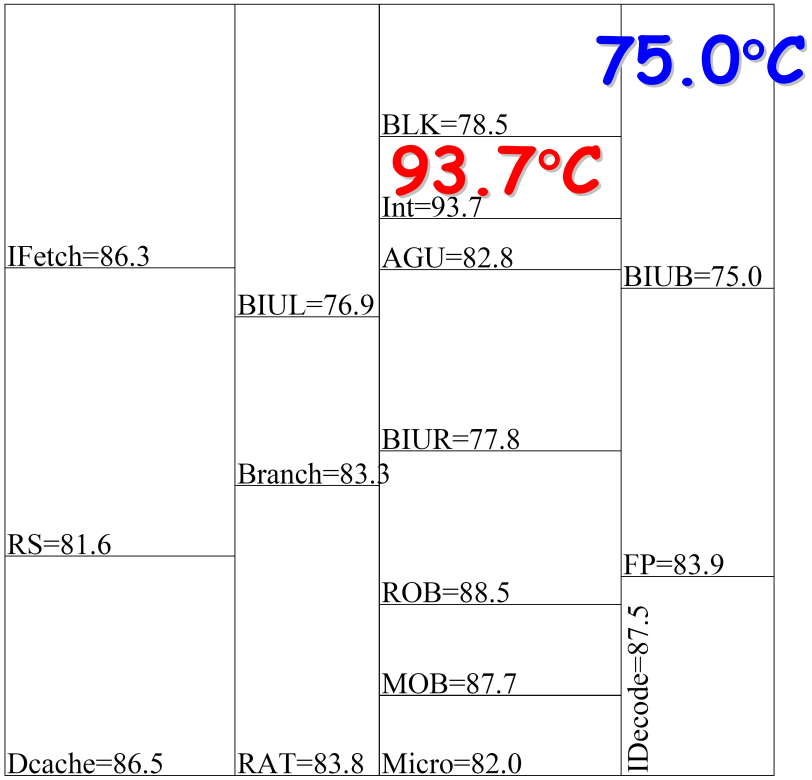
Pentium Pro

Original floorplan



Reduction in MaxTemp: 6.3°C
Increase in Wirelength: 13%

Low maximum temperature



Floorplan with highest maximum temperature 110.5°C

Smaller improvement due to smaller spread of temperature in Pentium Pro: 26°C vs 55°C in Alpha

Conclusions

- Our observations:
  - Maximum temperature reduction of 21°C while keeping a comparable wire length for Alpha processor
  - 6.3°C reduction in maximum temperature for Pentium Pro with a penalty of 13% in wire length
- Temperature-aware floorplanning can greatly improve temperature distribution of a chip