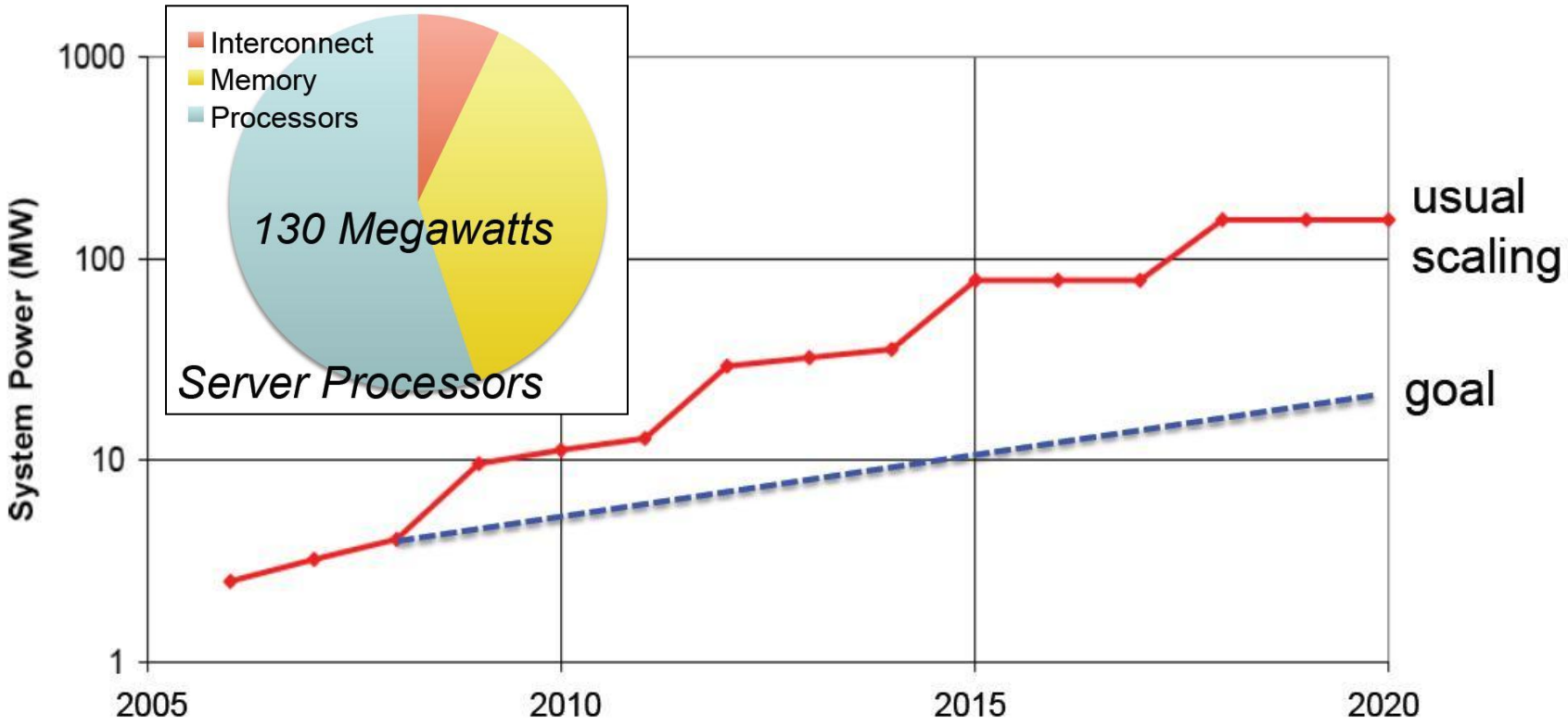

# Using Fast and Accurate Simulation to Explore Hardware/Software Trade-offs in the Multi-Core Era

WIM HEIRMAN, TREVOR E. CARLSON, SOURADIP SARKAR,
PIETER GHYSELS, WIM VANROOSE, LIEVEN EECKHOUT

*GHENT UNIVERSITY*
*UNIVERSITY OF ANTWERP*
*INTEL EXASCIENCE LAB*

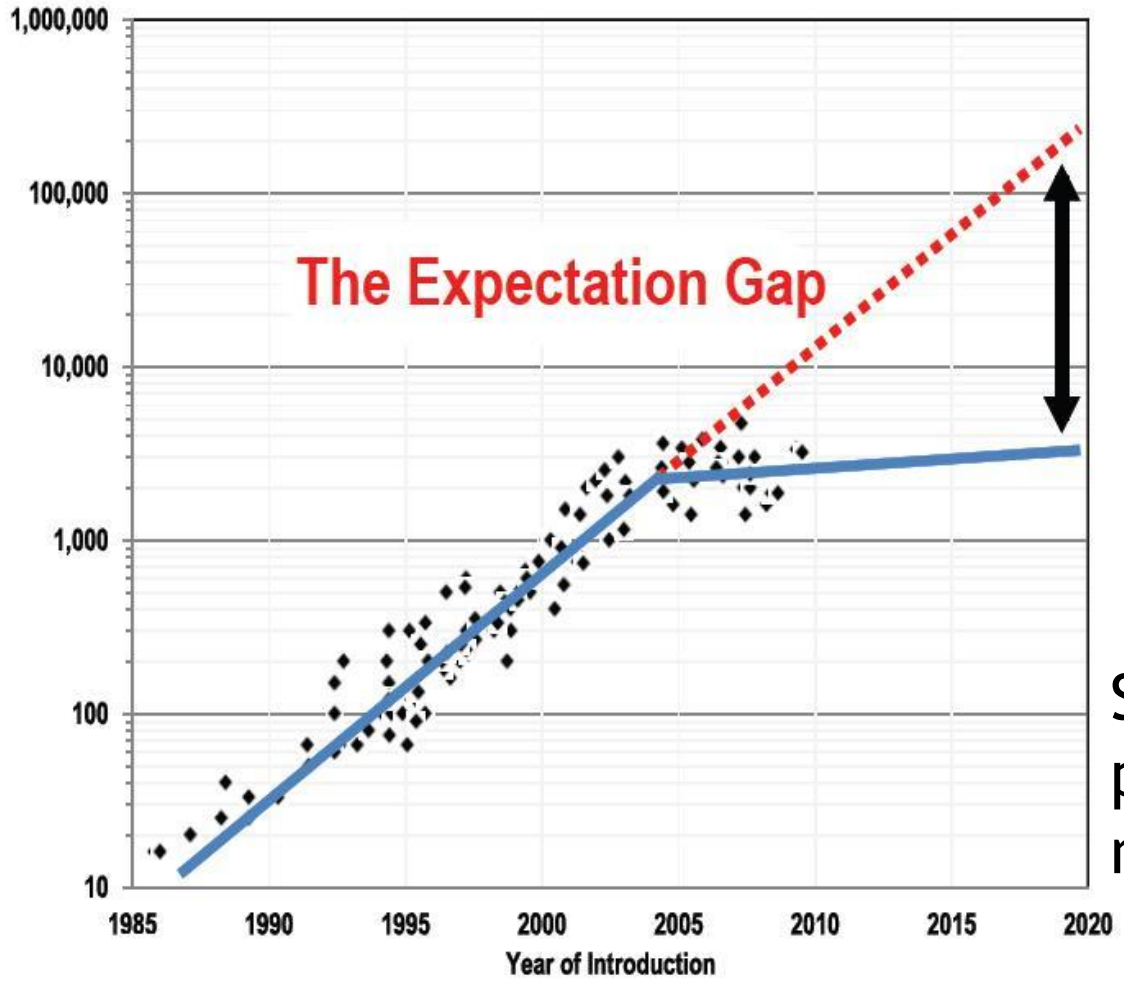
# HPC Power Challenges




- #1 on TOP500 (K,Japan) consumes 9.9MW @ 8.1 Petaflops
- Exascale goal: 2018, 20MW @ 1,000 Petaflops
  - 2x power, 100x performance (beyond Moore's law)

Source: Yalick, EXADAPT 2011

# HPC Performance Challenges



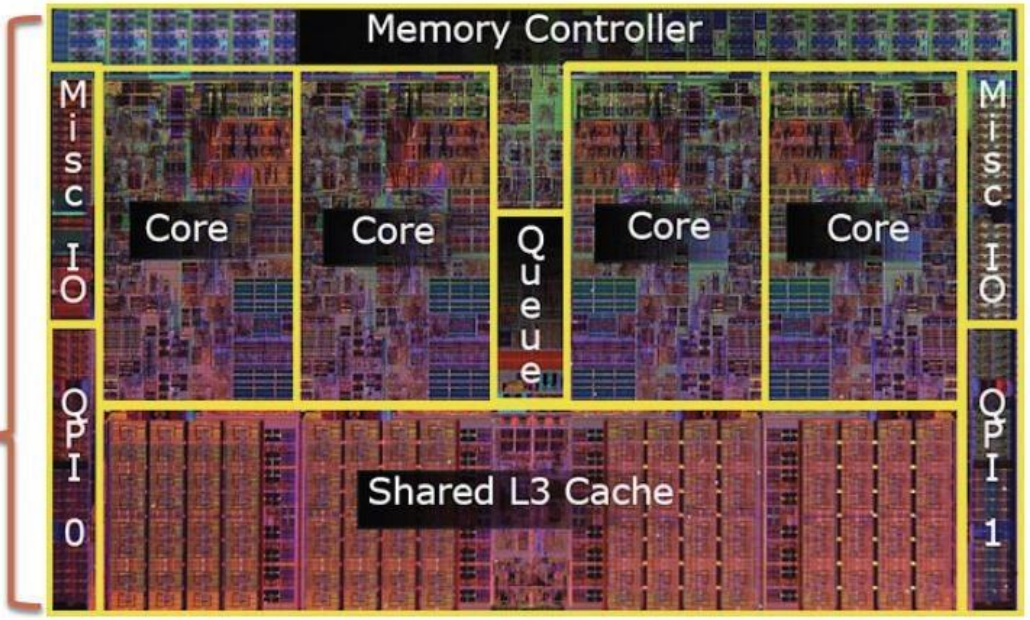

Single-core performance is not keeping pace

Source: Yalick, EXADAPT 2011

# Current Microprocessor Trade-offs

Cell phone processor
(0.1 Watt, 4 Gflop/s)

Server processor
(100 Watts, 50 Gflop/s)




Is there something in-between?

Source: Yalick, EXADAPT 2011

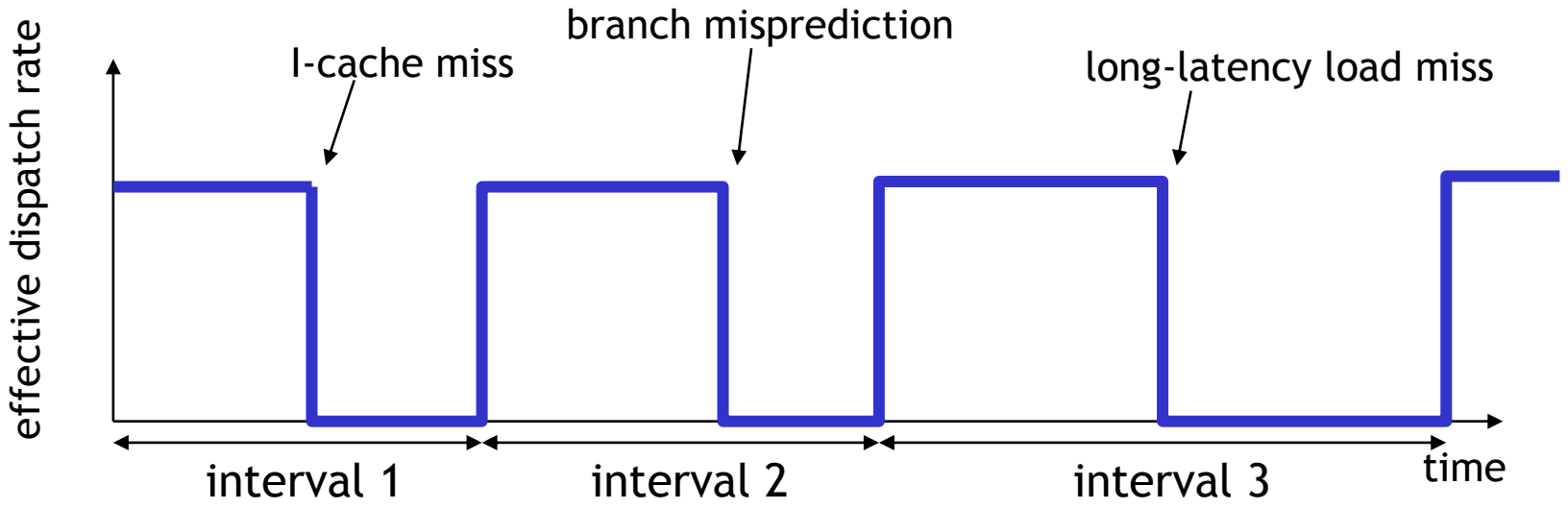
# Exascale Challenges

- Future systems will be diverse
  - Varying processor speeds
  - Varying failure rates for different components
  - Homogeneous applications become heterogeneous
- Software and hardware solutions are needed to solve these challenges
  - Handle heterogeneity (reactive load balancing)
  - Be fault tolerant
  - Improve power efficiency at the algorithmic level (extreme data locality)
- Hard to model accurately with analytical models

# Exascale challenges require fast and accurate simulation

- Fast and accurate simulation
  - Sniper = Interval Model + Graphite framework
  - Parallel simulator scales with the number of simulated cores
  - Available at http://www.snipersim.org
- Key Questions
  - What is the right level of abstraction?
  - When can we use these abstraction models?
- Key Capabilities
  - Model all (relevant) aspects of hardware accurately (enough)
  - Allow application designers to start exploring hardware-software interactions

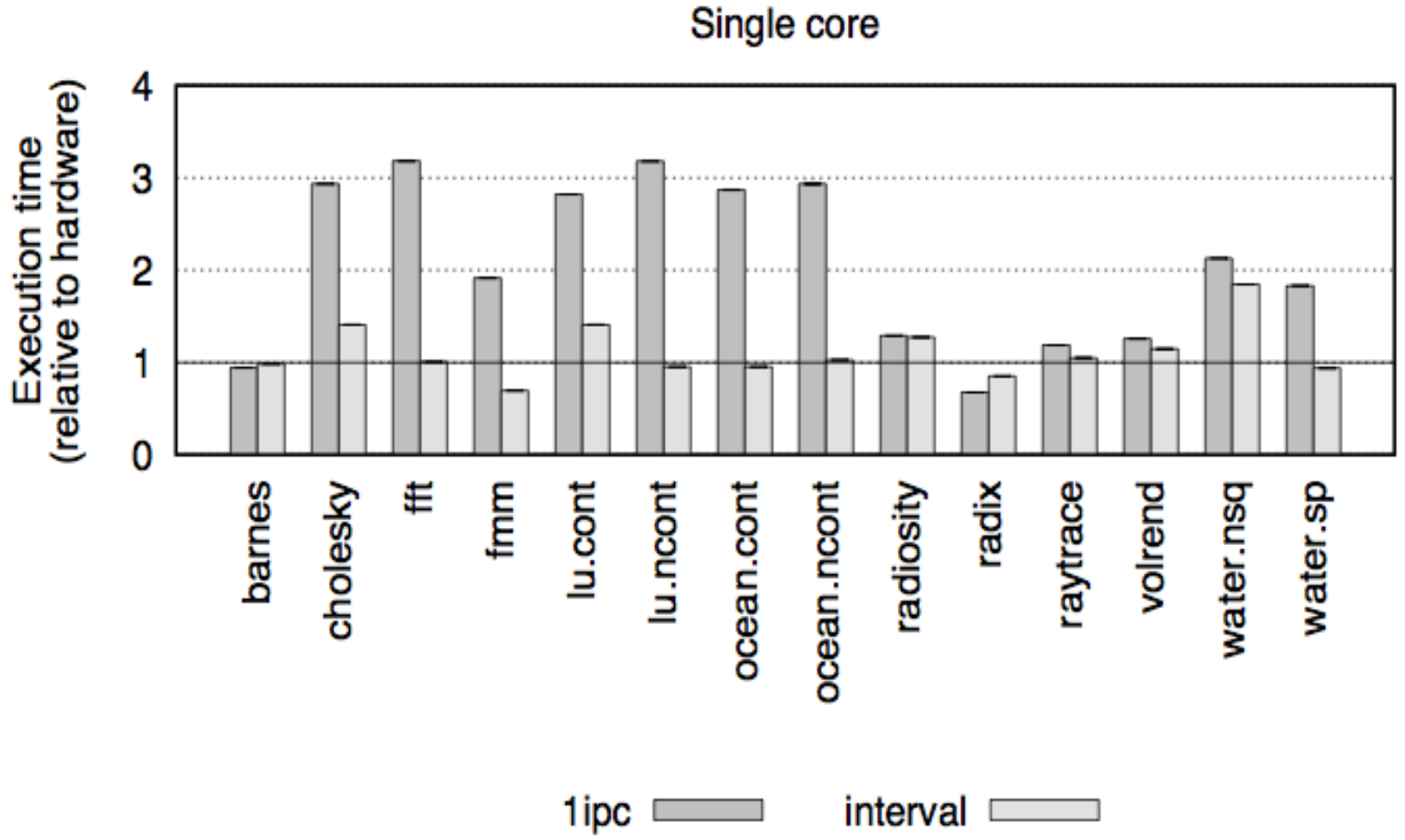
# Interval simulation: out-of-order core performance model at in-order simulation speed




*D. Genbrugge et al., HPCA'10*
*S. Eyerman et al., ACM TOCS, May 2009*
*T. Karkhanis and J. E. Smith, ISCA'04, ISCA'07*

# Simulator accuracy

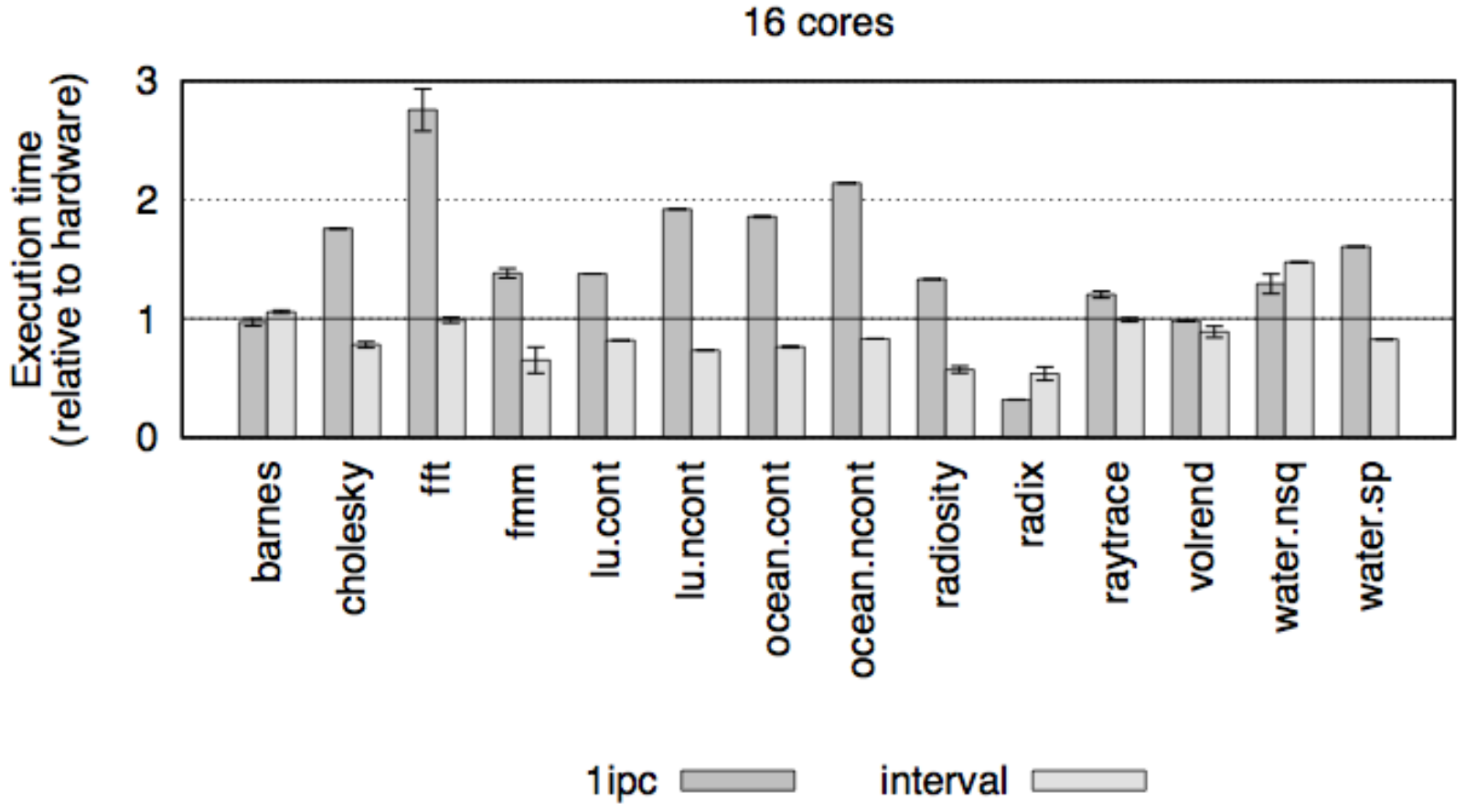
*single core: 20% avg abs error*




Source: Carlson, et. al, SC11, Nov 2011

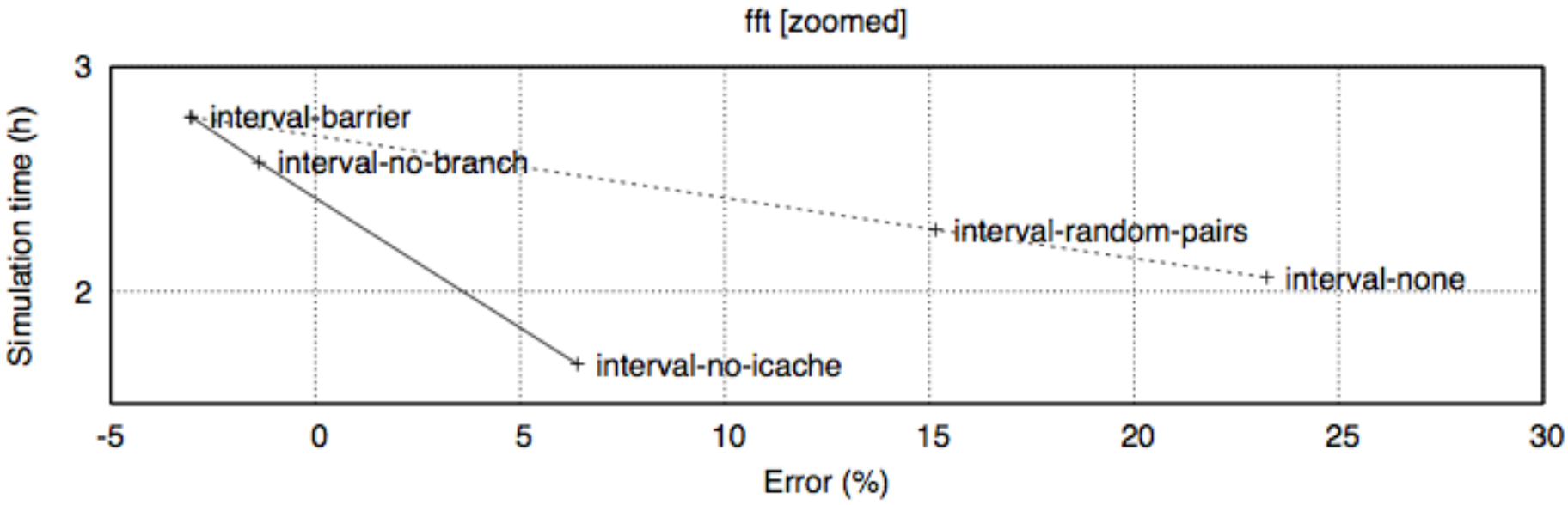
# Simulator accuracy

*16 cores: 24% avg abs error vs. 60% avg error for one-IPC*




Source: Carlson, et. al, SC11, Nov 2011

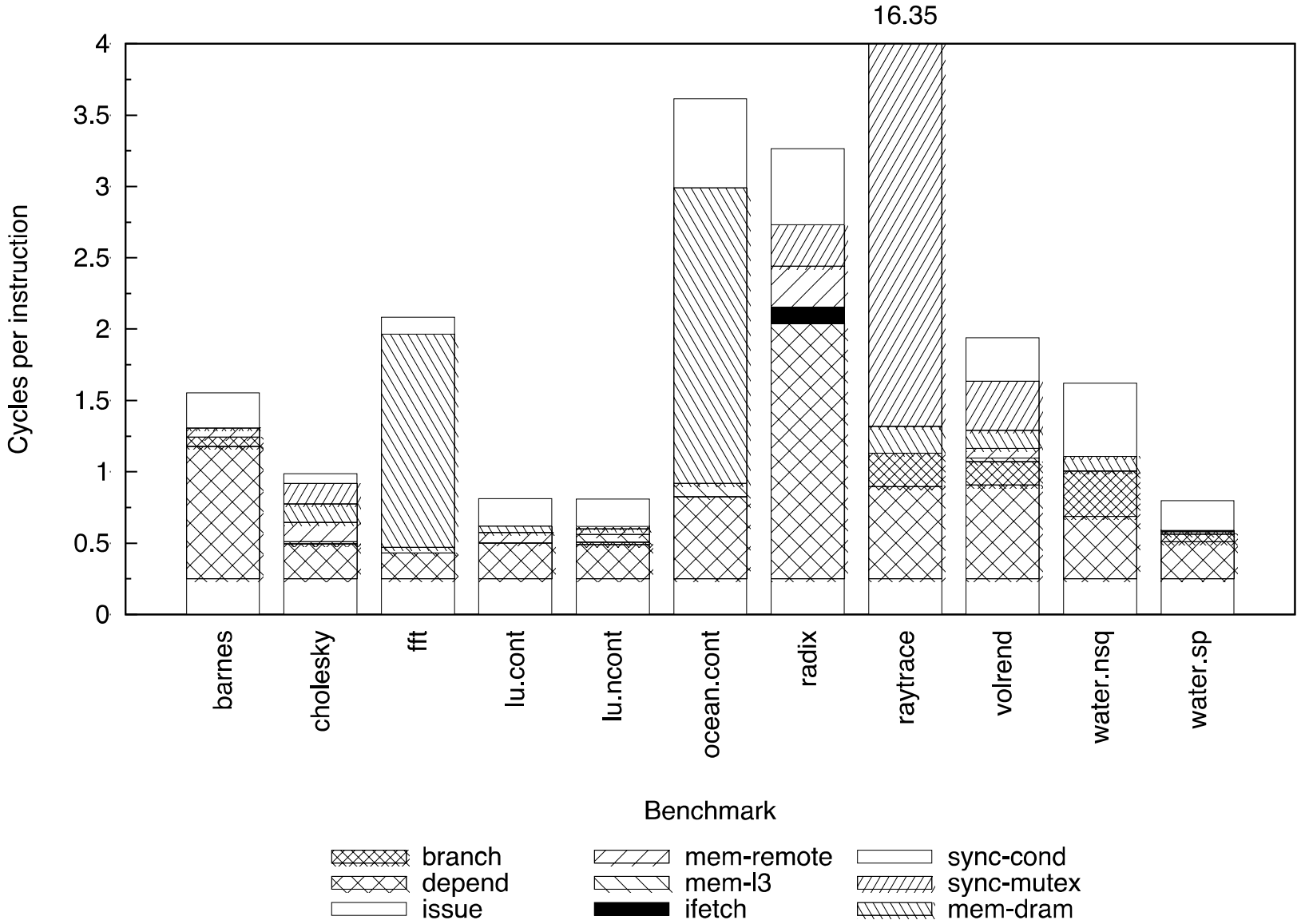
# Synchronization: speed vs. accuracy trade-off




Relaxed synchronization may compromise accuracy

Source: Carlson, et. al, SC11, Nov 2011

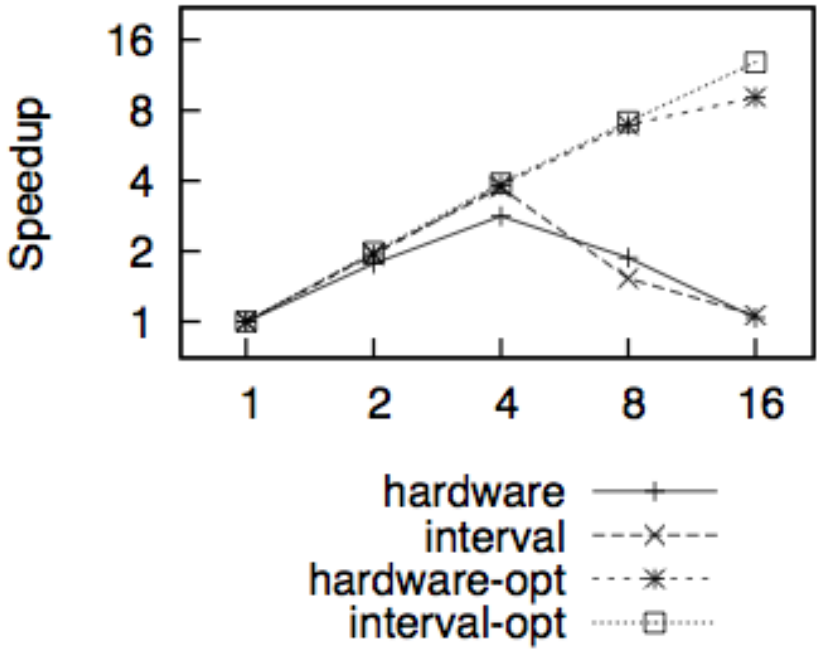
# Analyze performance bottlenecks

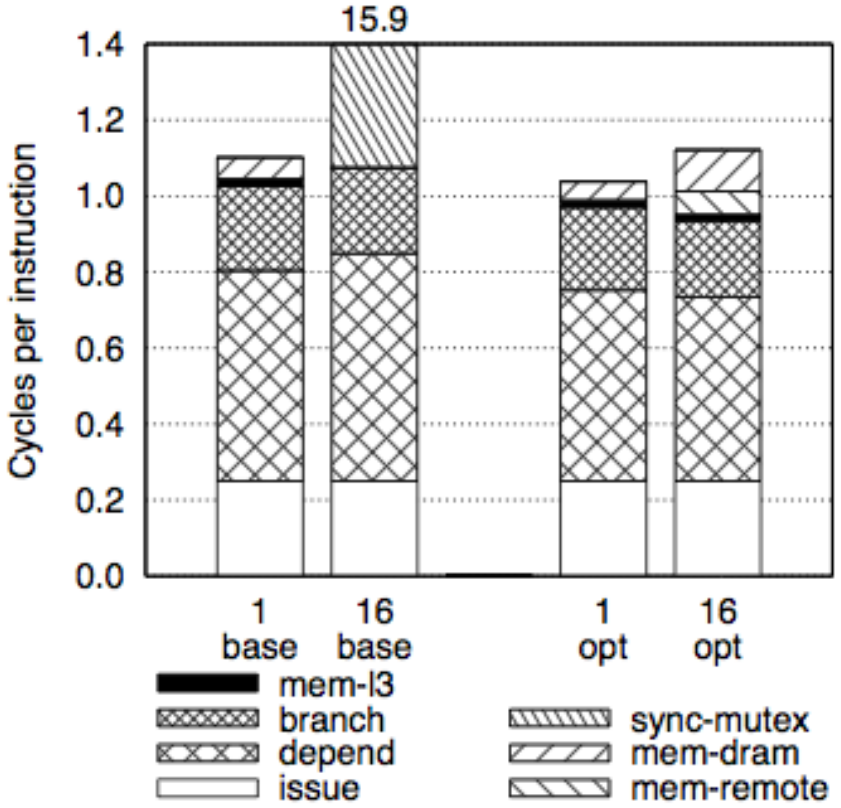



Source: Carlson, et. al, SC11, Nov 2011

# Suggest application improvements

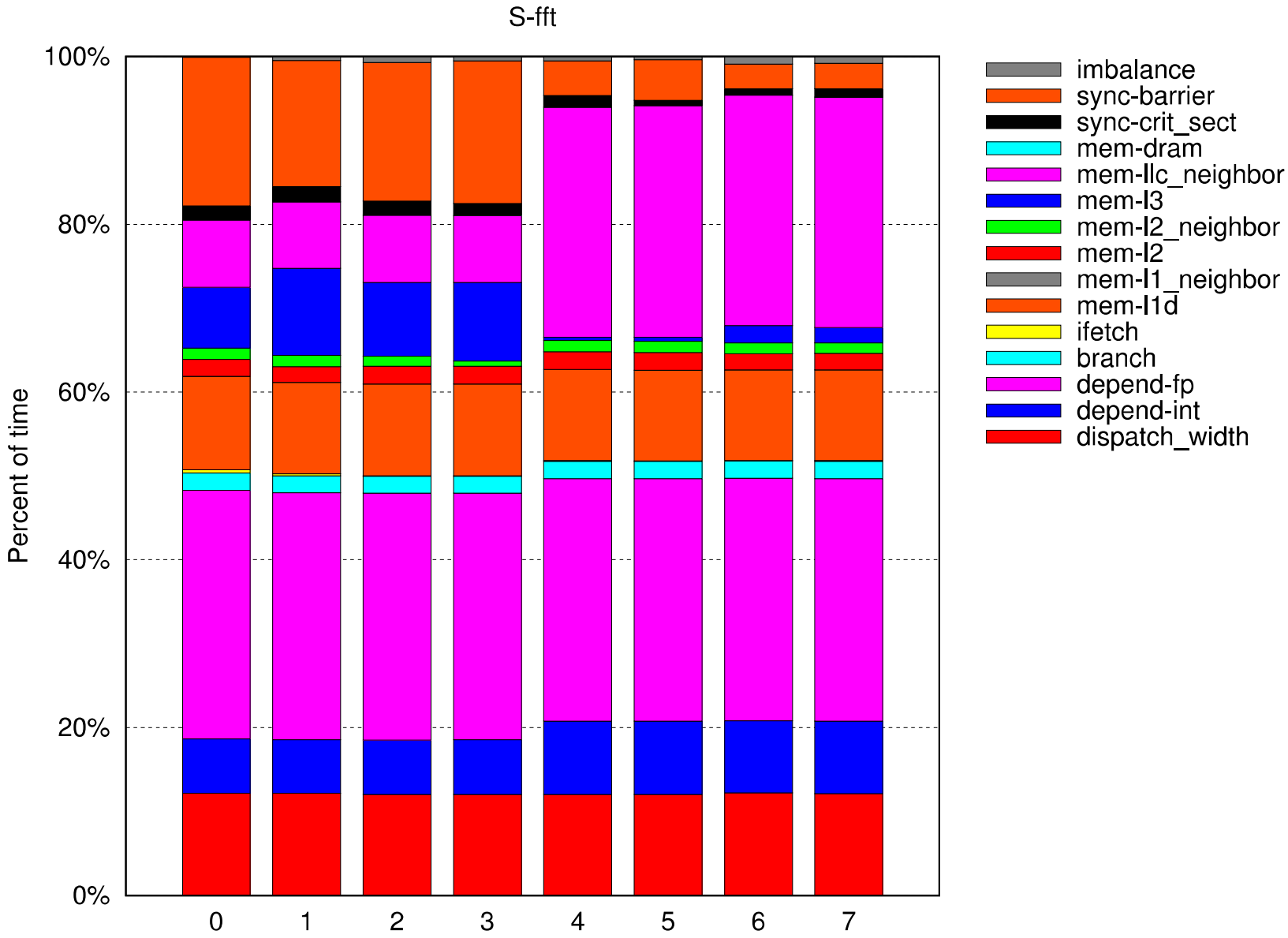
`pthread_mutex -> lock inc`




Source: Carlson, et. al, SC11, Nov 2011

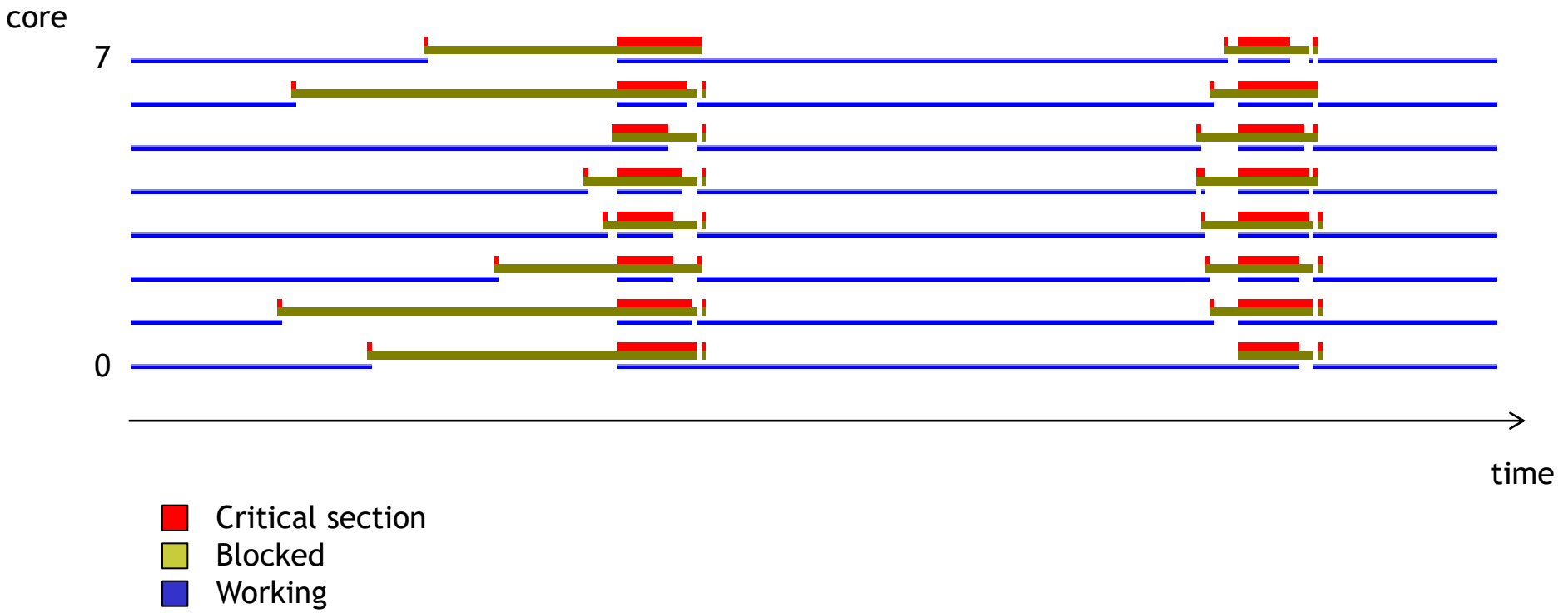
# Homogeneous applications?



# Analyze synchronization behavior

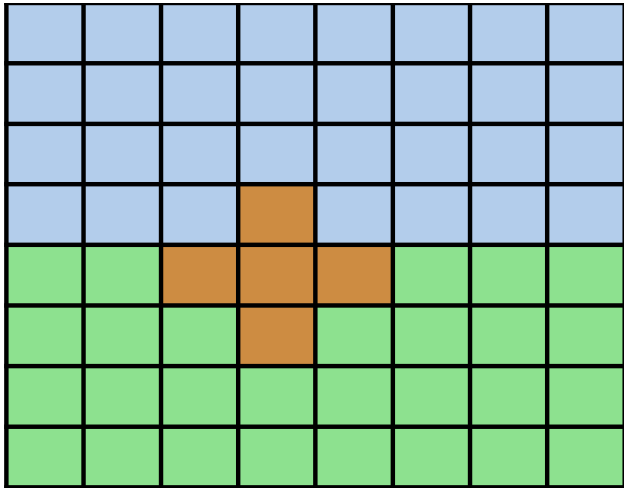
Thread-state timeline for *barnes*:

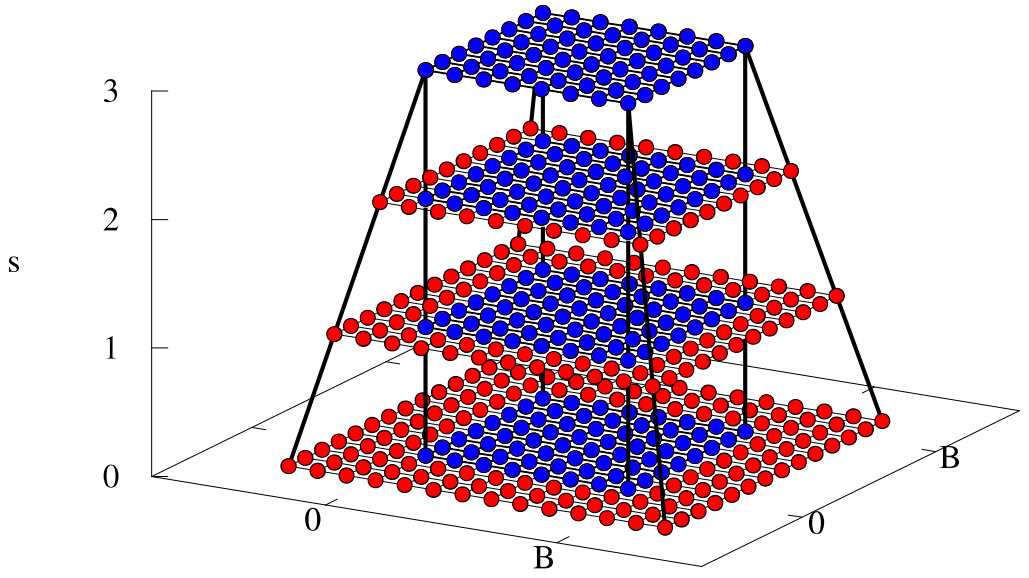
synchronization, critical sections, load imbalance



# Case study: tiled heat transfer

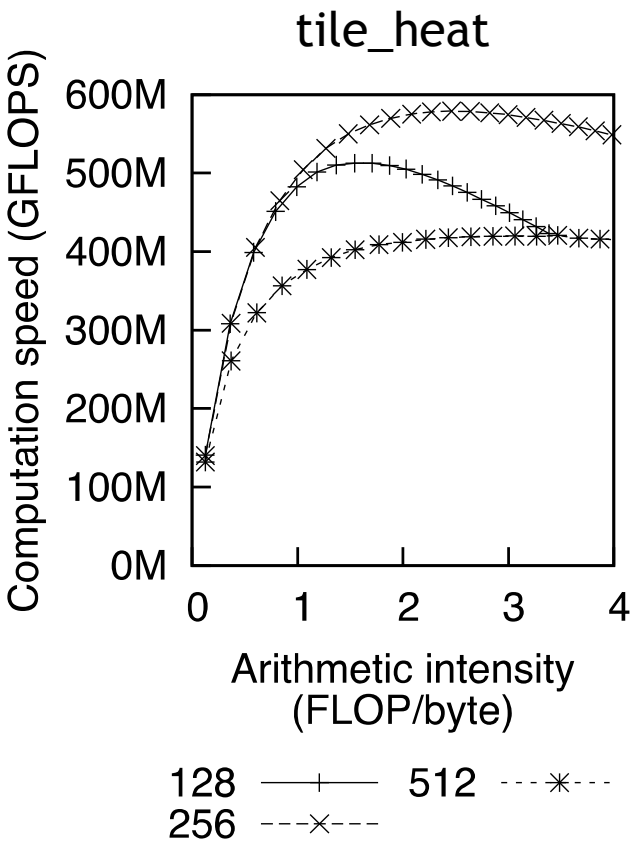
- 5-point stencil, applied to consecutive time steps
- Optimization: tiled to optimize locality, multiple time steps per tile – but requires redundant computation at tile edges

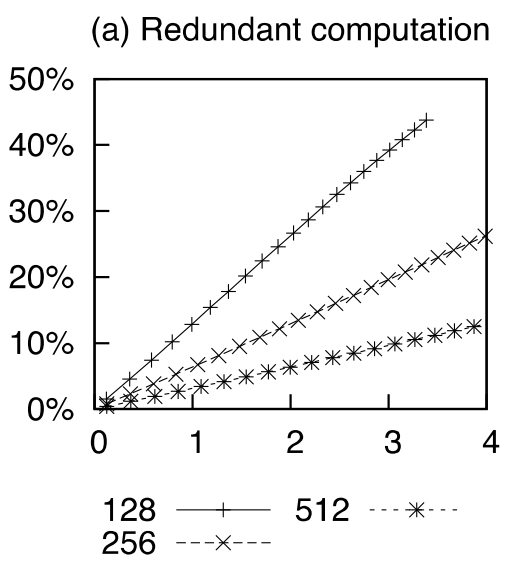


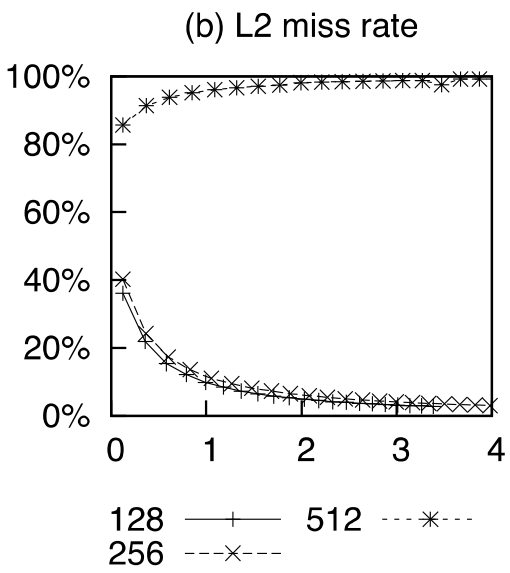



Source: P. Ghysels, 2011

# Tile size and steps vs. cache behavior



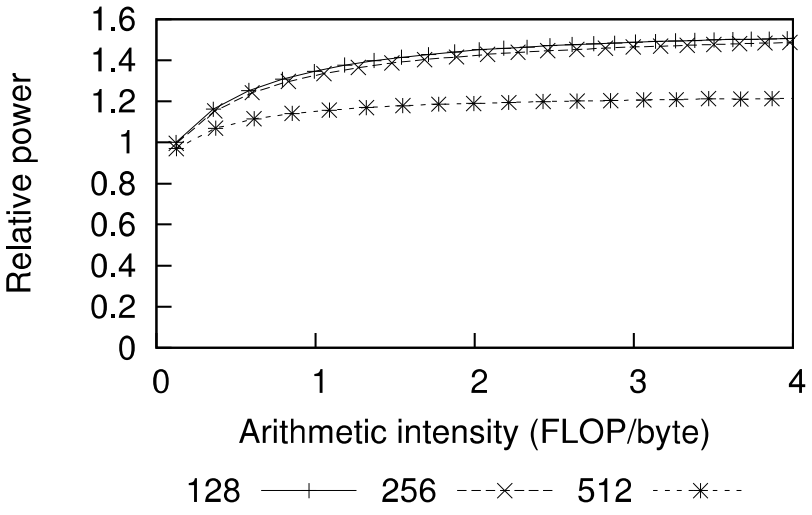



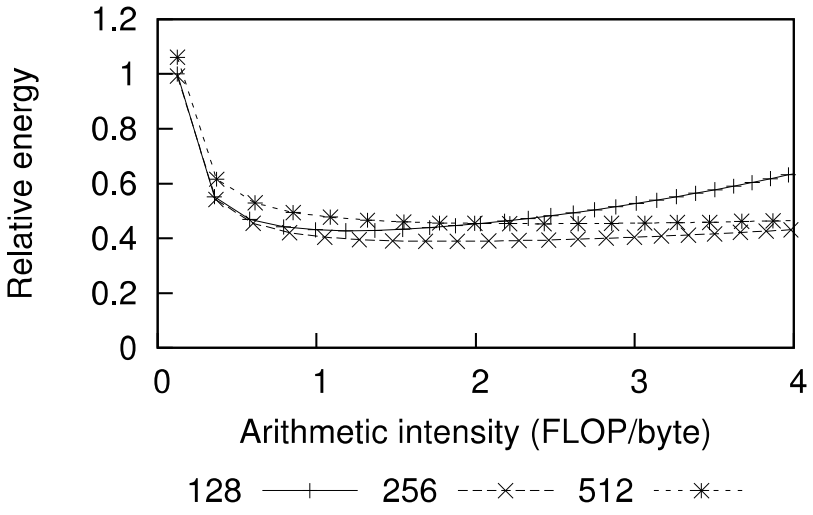
# Not just time, energy as well

Integration with McPAT, provide application-specific estimates for power, energy (and EDP, $ED^2P$, …)

(a) Power

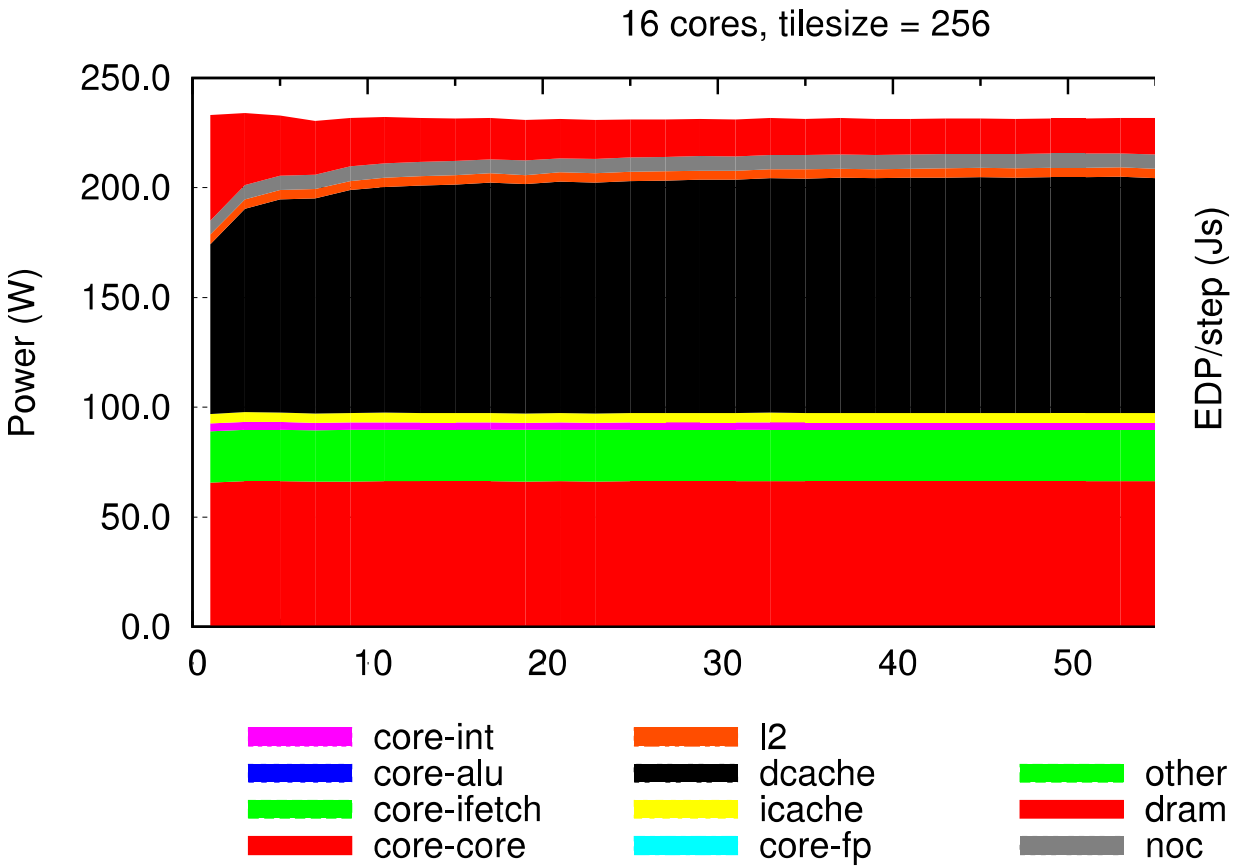


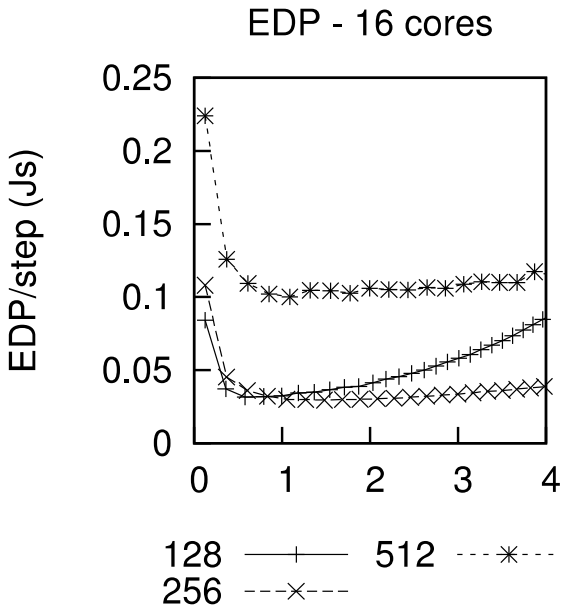

(b) Energy per step




*McPAT: Li et al., MICRO'09*

# Where is the energy going?



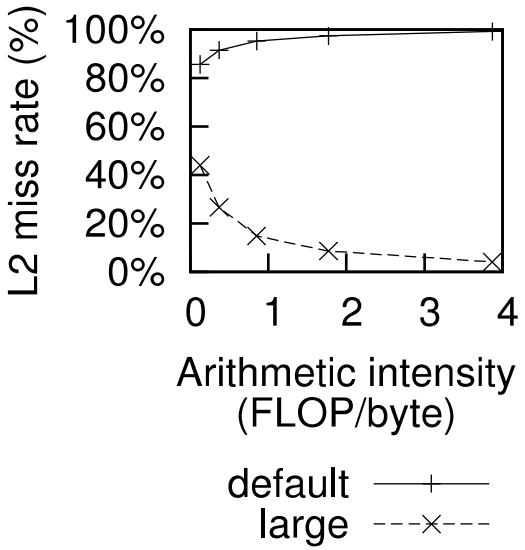




EDP: Energy Delay Product

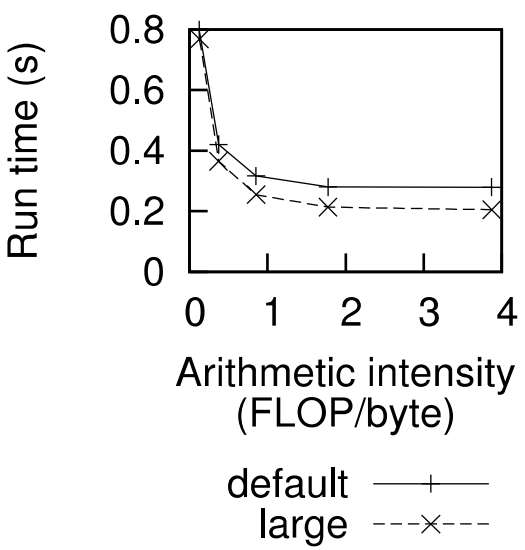
# Architectural exploration

Experiment: double the size of L2 and L3 caches

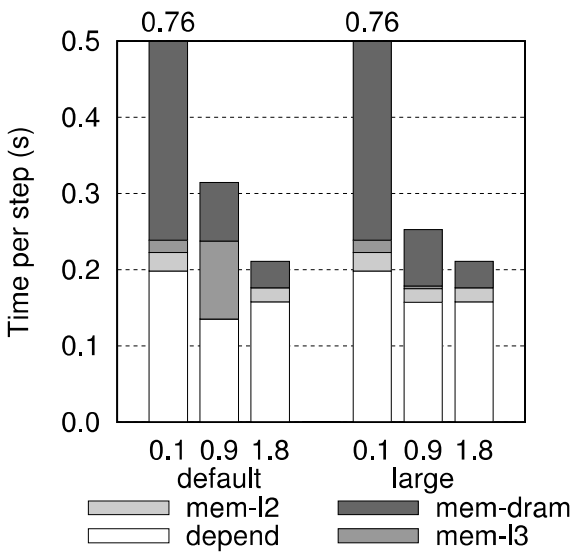
(a) L2 cache miss rate




(b) Run time per step




(c) Time stack



# Conclusions

- Sniper is a fast and accurate simulation for multicore processors
- Useful both for hardware and software designers
- Faster than most simulators, so Sniper can be used to model the effects of large caches, large input sets, multiple runs
- More accurate than instrumentation (no intrusion), higher visibility
- Allows for architectural exploration (vs. performance counters)


# Using Fast and Accurate Simulation to Explore Hardware/Software Trade-offs in the Multi-Core Era

WIM HEIRMAN, TREVOR E. CARLSON, SOURADIP SARKAR,
PIETER GHYSELS, WIM VANROOSE, LIEVEN EECKHOUT

*GHENT UNIVERSITY*
*UNIVERSITY OF ANTWERP*
*INTEL EXASCIENCE LAB*