Project
Algorithm Engineering

Jens K. Mueller

jkm@informatik.uni-jena.de

Department of Mathematics and Computer Science
Friedrich-Schiller-University Jena
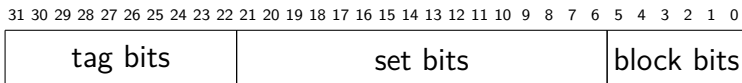
Wednesday 23rd April, 2014

Caches

Memory Hierarchy

Cache

- Sets (S), lines (E), and blocks (B)
  The capacity of a cache is

$$C = S \cdot E \cdot B$$

- Split memory addresses into

31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

tag bits	set bits	block bits

How to obtain information about caches?

Lookup

1. Choose set given the set bits
2. Find valid line with given tag bits (hit or miss)
3. If hit, return block given block bits
4. Else fetch memory and place into cache

Cache Misses

- Cold miss
- Conflict miss
- Capacity miss

$E = ?$

- Direct-mapped ($E = 1$)
- n-way associative ($E = n$)
- Fully-associative ($S = 1$, i.e. $E = C/B$)

Line Replacement Strategies

- Random
- Least Recently Used
- Least Frequently Used

Write Strategy

Write Hits
- ▶ Write through
- ▶ Write back (needs dirty bit)

Write Misses
- ▶ Write allocate
- ▶ No write allocate

Assume write back and write allocate!

Cache Performance

Cache Performance Metrics

- ▶ Miss rate (#misses/#references)
- ▶ Hit rate (1 - miss rate)
- ▶ Hit time
- ▶ Miss penalty time

Trade-Offs

- ▶ Cache size
- ▶ Line size
- ▶ Associativity
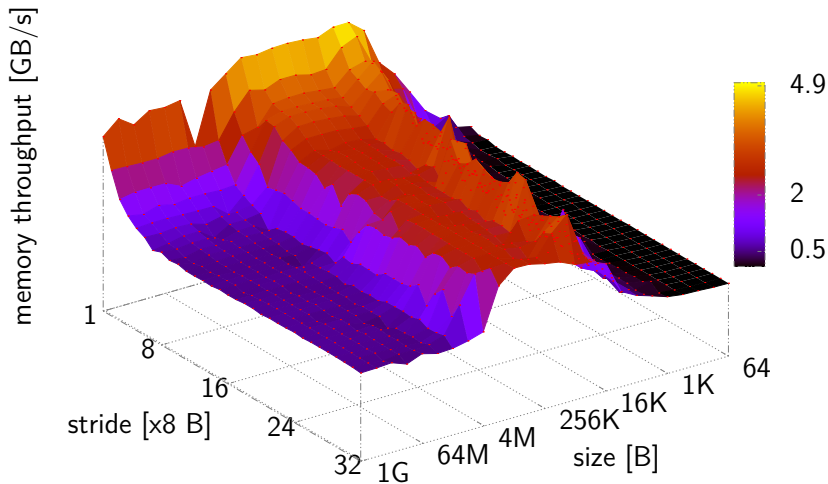- ▶ Write Strategy

Cache-Friendly Code

- Locality (spatial/temporal) tends to lower miss rates
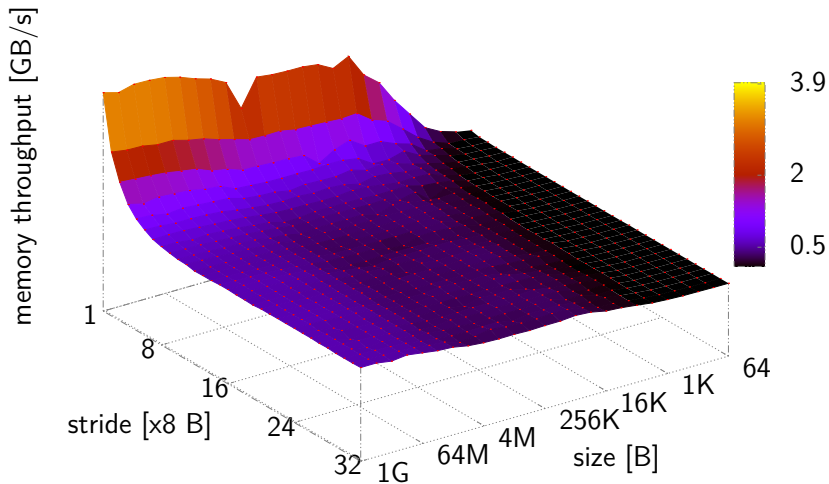- Minimize cache misses

Measuring Cache Misses

STREAM benchmark

John D. McCalpin. STREAM: Sustainable Memory Bandwidth
in High Performance Computers. URL:
http://www.cs.virginia.edu/stream/

Memory Mountain

Memory Mountain

Homework

- ▶ Plot the memory mountain for your system
- ▶ Measure the memory throughput for your project

References

📄 Randal E. Bryant and David R. O'Hallaron.
Computer Systems: A Programmer's Perspective.
2nd. USA: Addison-Wesley, 2010. ISBN:
0136108040, 9780136108047.

📄 John D. McCalpin. STREAM: Sustainable Memory
Bandwidth in High Performance Computers. URL:
http://www.cs.virginia.edu/stream/ (cit. on
p. 14).