

# N-body memory layout exploration

Oliver Geisel & Lisa Hentschke

January 30, 2021

# Structure

The task

Solution Strategies

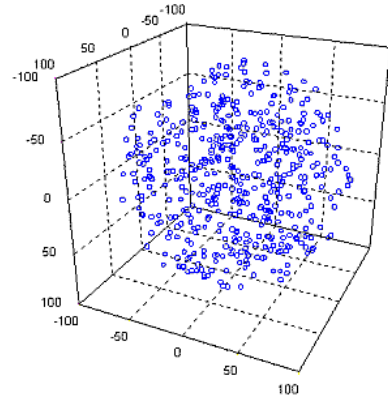
Results

Explanation

Further Approaches

# The n-body simulation

- ▶ simulate the interaction of  $n$  particles
- ▶ each particle has
  - ▶ position x
  - ▶ position y
  - ▶ position z
  - ▶ velocity x
  - ▶ velocity y
  - ▶ velocity z
  - ▶ mass



<http://astro.dur.ac.uk/~nm/pubhtml/nbody/nbody.html>

# Solution Strategies

- ▶ rewrite CPP-code in CUDA: implement AoS, SoA and AoSoA memory layouts
- ▶ implemented shared memory variants
- ▶ for SoA: implemented two sub-variants: K and T
  - ▶ B: compute one particle per block
  - ▶ T: compute one particle per thread

# Example

We tested on K80 (Taurus), v100 (Taurus), and 1070 (private, driver version 461.09) respectively

TESTED ON  
Tesla K80

---

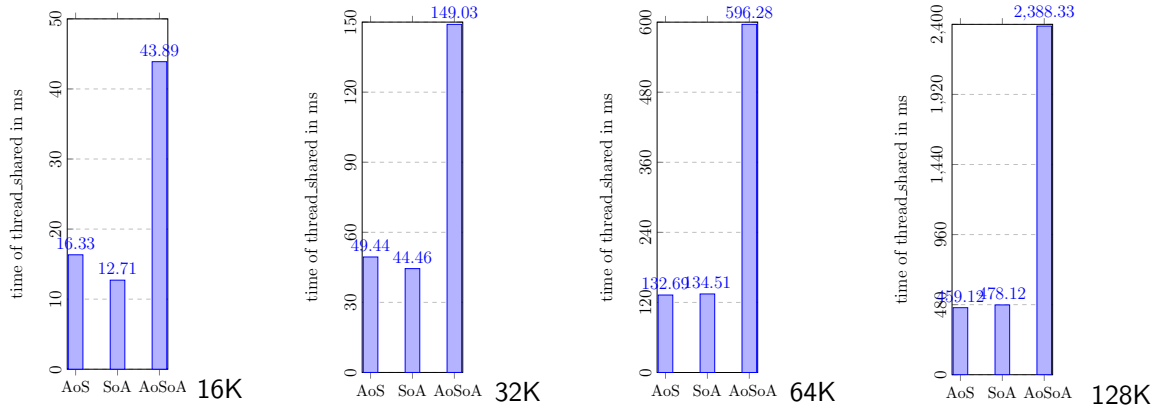
16 k particles (448.000 kiB)

Benchmarks:	Thread,	Thread_shared,	Move
AoS	56.5121ms	17.3868ms	0.036544ms
AoS	56.4797ms	17.4133ms	0.033280ms
AoS	54.9357ms	15.6404ms	0.033632ms
AoS	50.7257ms	15.6060ms	0.032544ms
AoS	50.7504ms	15.5858ms	0.032768ms
AVG:	53.8807ms	16.3264ms	0.033754ms

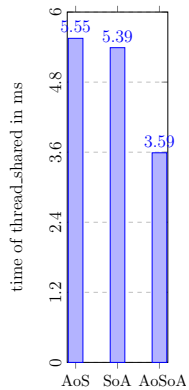
Benchmarks: Block,	Block_shared,	Thread,	Thread_shared,	Move
SoA 55.0906ms	21.4492ms	22.8277ms	14.2324ms	0.0110ms
SoA 54.3634ms	21.4436ms	22.8197ms	13.0962ms	0.0083ms
SoA 49.3135ms	19.3027ms	20.4029ms	12.7189ms	0.0083ms
SoA 48.4277ms	18.4641ms	18.8272ms	11.7521ms	0.0092ms
SoA 44.3154ms	17.7516ms	18.8140ms	11.7349ms	0.0092ms
AVG: 50.3021ms	19.6823ms	20.7383ms	12.7069ms	0.009184ms

Benchmarks:	Thread,	Thread_shared,	Move
AoSoA	28.2286ms	43.9762ms	0.222112ms
AoSoA	28.2296ms	43.6562ms	0.220672ms
AoSoA	28.2461ms	43.9120ms	0.221504ms
AoSoA	28.2501ms	43.8764ms	0.222688ms

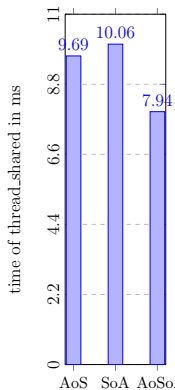
# Compare the memory structures - on K80



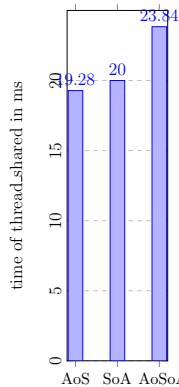
# Compare the memory structures - on v100



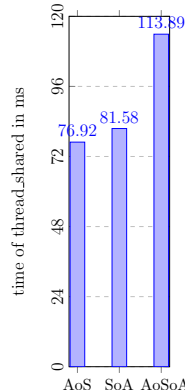
16K



32K

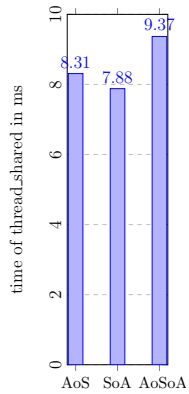


64K

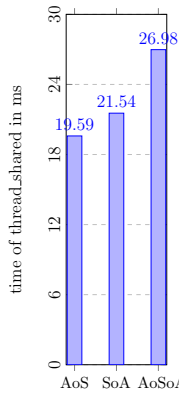


128K

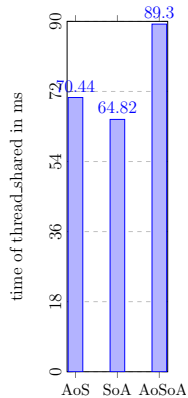
# Compare the memory structures - on 1070



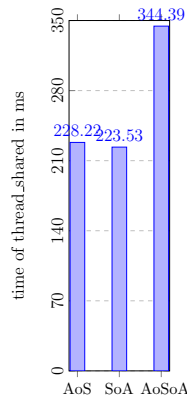
16K



32K



64K



128K



# Why?

Memory Layout:

- ▶ K80 - GDDR 5 with SDRAM
- ▶ v100 - HBM 2
- ▶ 1070 - GDDR 5

# What else could we try?

- ▶ use shared (we did that)
- ▶ use texture memory
- ▶ change computation (not part of the task)