Approach to
Data-Oriented Design

Omid Shahbazi
https://omidshahbazi.github.io

Agenda

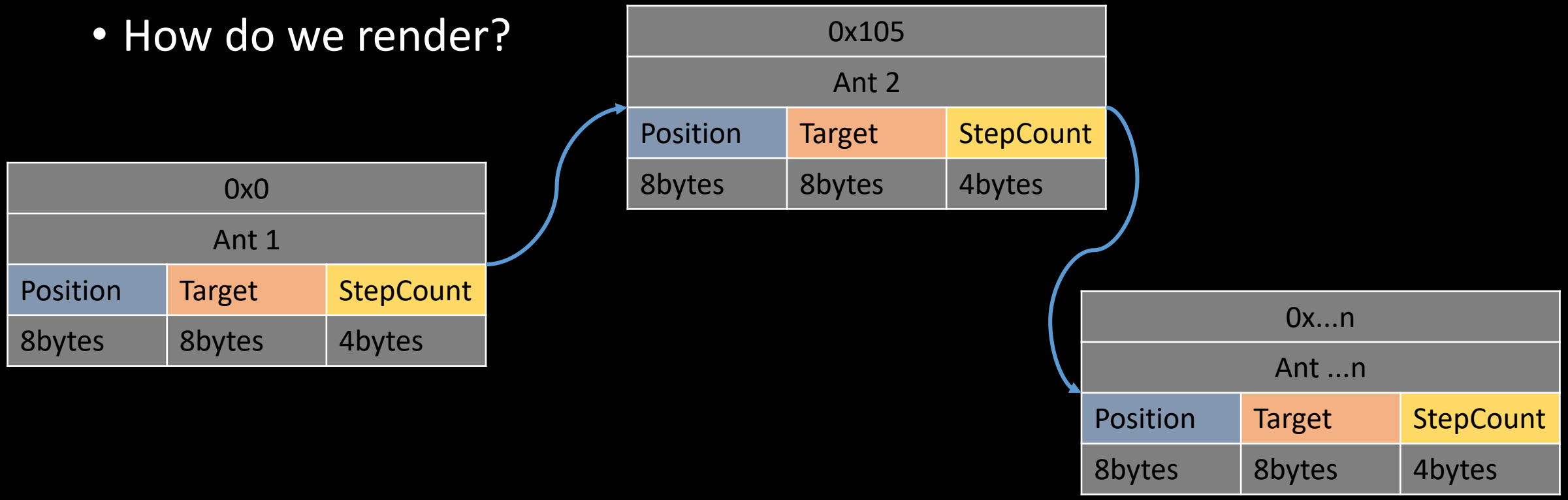
- Problem Statement
- Discussion About Possible Solutions
- What's the actual problem?
- Go Over the Problem
- Studing Better Impl.
- Q/A

Problem Statement

- Simulate behavior of, and Render massive ants crawling on a map
  - Each ant is a pixel
  - We’re going to benchmark and impl. some solutions
- Do you need more to know?

Let’s Discuss

- What is our data?
- What is our behavior(s)?
- How do we render?



Let’s see the impl.

- How long should it take?
- Give me some solutions

Optimize

- Are we happy?
- Any other solution?

Optimize

- Micro optimization !?
- Change the memory layout !?
- Anything else ?

0x0			0x20			0x...n x 20		
Ant 1			Ant 2			Ant ...n		
Position	Step	StepCount	Position	Step	StepCount	Position	Step	StepCount
8bytes	8bytes	4bytes	8bytes	8bytes	4bytes	8bytes	8bytes	4bytes

What’s the actual problem?

- Cost of Abstractions
- Branch (mis)predictions
- Cost of Hitting RAM
- Cache Misses
- So, what we have to do?

Benchmark Memory Access

ElementSize: 4b Count: ~269M Size: 1gb CPU-Freq.: ~4.3GHz

RandomMemoryAccess is Running

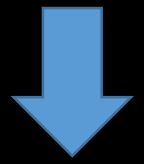
Cycles: ~17M Time: ~4.3s Speed: ~242mb/s

SequentialMemoryAccess is Running

Cycles: ~1M Time: ~327ms Speed: ~3.2gb/s

Studing better impl.

0x0			0x20			0x...n x 20		
Ant 1			Ant 2			Ant ...n		
Position	Step	StepCount	Position	Step	StepCount	Position	Step	StepCount
8bytes	8bytes	4bytes	8bytes	8bytes	4bytes	8bytes	8bytes	4bytes



8bytes	Position	Position	Position	Position	Position	Position
12bytes	Step	StepCount	Step	StepCount	Step	StepCount

Data-Oriented Design Principles

- If you don’t understand the data, you don’t understand the problem.
- Different problems, require different solutions.
- If you have different data, you have different problem.
- If you don’t understand the cost of solving the problem, you don’t understand the problem.
- If you don’t understand the hardware, language, compiler, OS you can’t reason about the cost of solving the problem.
- It’s just a name, understand the hardware, language, compiler, OS, data, then solve the problem

Thank You!

- References
  - Anger Fog, Optimization Manuals
    - https://agner.org/optimize
  - Richard Fabian, “Data-Oriented Design”
    - http://dataorienteddesign.com/dodbook
  - Yaser Zhian, “Data Oriented”
    - http://yzt.github.io/
- Any Questions?

https://omidshahbazi.github.io