

Team E1 - FPGA Accelerated Fluid Simulation

Jeremy Dropkin, Alice Lai, Ziyi Zuo

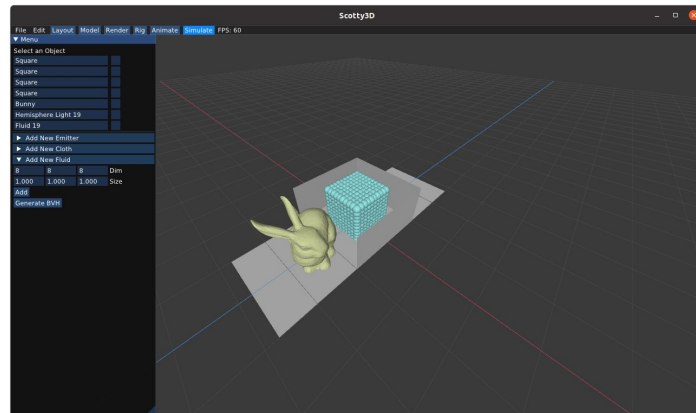
Add your 12 slides after this slide... [remember, 12 min talk + 3 min Q/A]



Motivation

— — —

- Scotty3D is a 3D graphics software package that runs only as a multithreaded CPU Program
- Fluid Simulation Library runs poorly on CPU, throttling the FPS to fractional values
- Our goal is to create an **FPGA-based application that can accelerate Scotty3D's Fluid Simulation Library**

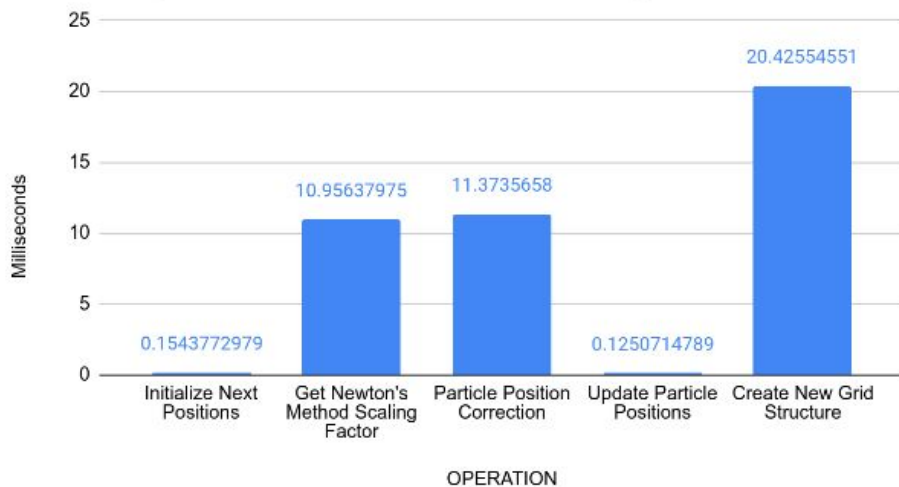


Scotty3D Interface

Performance Requirements

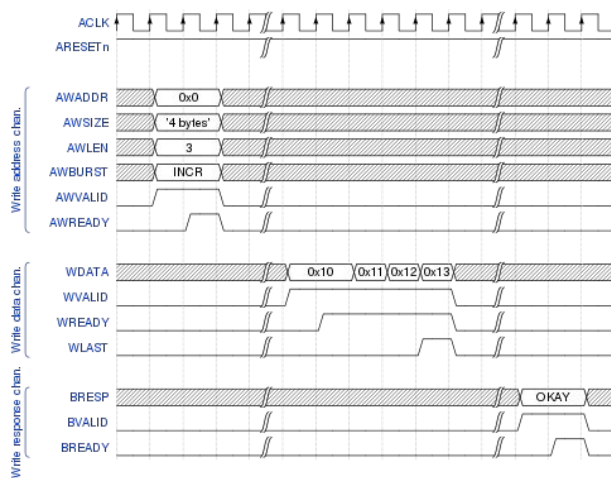
- Baseline
 - ~3 second render time (per frame) on an i7-8665U @ 1.90 GHz x 8
- Our goal is to make at least a **10x speedup**
 - ~300ms to render per frame

Timing Breakdown of Fluid Simulation Update Iteration



Data Communication Requirements

- Data transfer between CPU and fabric should not bottleneck the actual rendering task
- Minimize interpretation overhead between CPU and FPGA format
 - CPU/Fabric communication use AXI interface



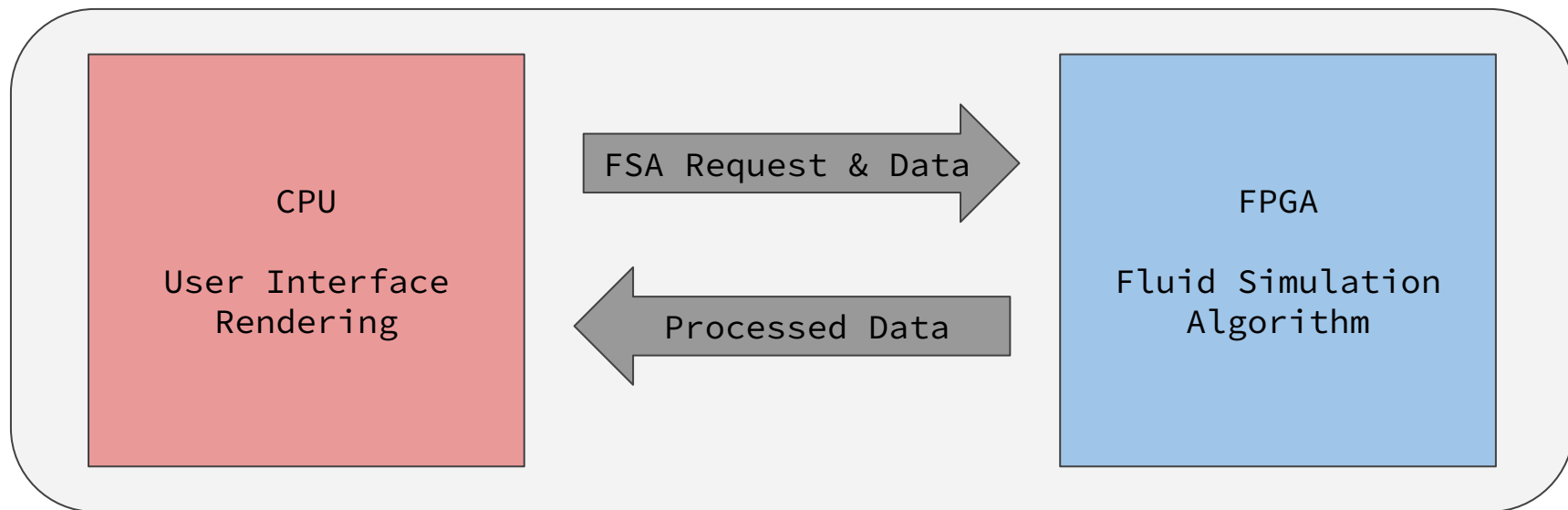
AXI Interface Waveform Diagram

Use Case Requirements

- Accelerated implementation does not add to overhead or complications to the user experience
 - We will collect user feedback through user testing to ensure there is no additional overhead
- Perform computation within a 10W TBP envelope
 - Ensure low-power compute
 - Desktops typically run between 50W and 100W

Solution Approach

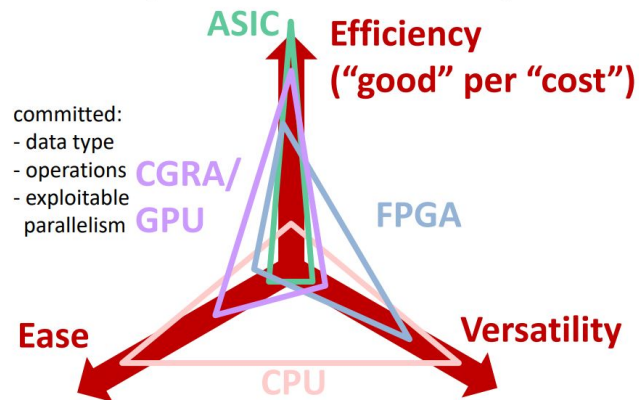
- Accelerate Fluid Simulation Algorithm on the FPGA fabric
- CPU handles the rest of the Scotty3D rendering stack



Solution Platform - Motivation

Why FPGA?

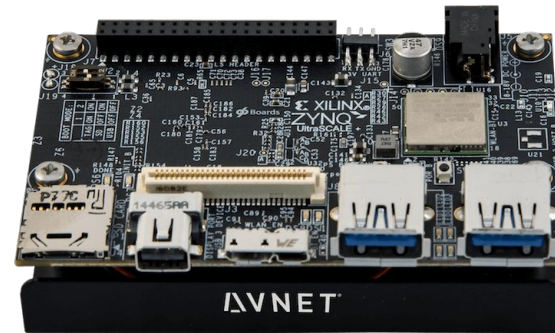
- CPU multithreading is not enough to take advantage of high thread-level parallelism
- GPUs are capable of exploiting thread-level parallelism using vectorized operators
- FPGA!
 - Extract high thread-level parallelism
 - Handle irregular code by instantiating extra control structures
 - Highly flexible memory architectures



From 18643 course slides

Solution Platform

- Xilinx Vivado Design Suite
 - Development and synthesis platform for Xilinx FPGA platforms
- Vitis HLS
 - High Level Synthesis
 - Generate hardware from C/C++ code
 - Control structures through use of pragmas and compiler directives
 - Result is transpiled to HDL
- Xilinx Ultra96 FPGA
 - Arm Core + Reconfigurable Fabric



Solution Approach - Algorithm

- Exploit thread-level parallelism
 - Unroll loops and instantiate multiple computational units
- Exploit FPGA BRAMS
 - Block RAM - high speed versatile memories
 - Large BRAM capacity means we can make multiple copies of data
 - Use this to accelerate neighbor-finding tasks

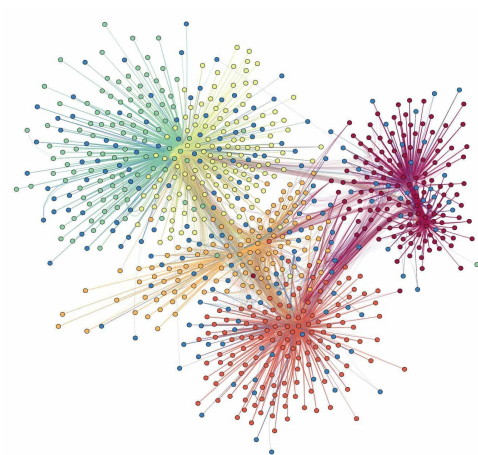
Algorithm 1 Simulation Loop

```
1: for all particles  $i$  do
2:   apply forces  $\mathbf{v}_i \leftarrow \mathbf{v}_i + \Delta t \mathbf{f}_{ext}(\mathbf{x}_i)$ 
3:   predict position  $\mathbf{x}_i^* \leftarrow \mathbf{x}_i + \Delta t \mathbf{v}_i$ 
4: end for
5: for all particles  $i$  do
6:   find neighboring particles  $N_i(\mathbf{x}_i^*)$ 
7: end for
8: while  $iter < solverIterations$  do
9:   for all particles  $i$  do
10:    calculate  $\lambda_i$ 
11:   end for
12:   for all particles  $i$  do
13:    calculate  $\Delta \mathbf{p}_i$ 
14:    perform collision detection and response
15:   end for
16:   for all particles  $i$  do
17:    update position  $\mathbf{x}_i^* \leftarrow \mathbf{x}_i^* + \Delta \mathbf{p}_i$ 
18:   end for
19: end while
20: for all particles  $i$  do
21:   update velocity  $\mathbf{v}_i \leftarrow \frac{1}{\Delta t} (\mathbf{x}_i^* - \mathbf{x}_i)$ 
22:   apply vorticity confinement and XSPH viscosity
23:   update position  $\mathbf{x}_i \leftarrow \mathbf{x}_i^*$ 
24: end for
```

Technical Challenges

— — —

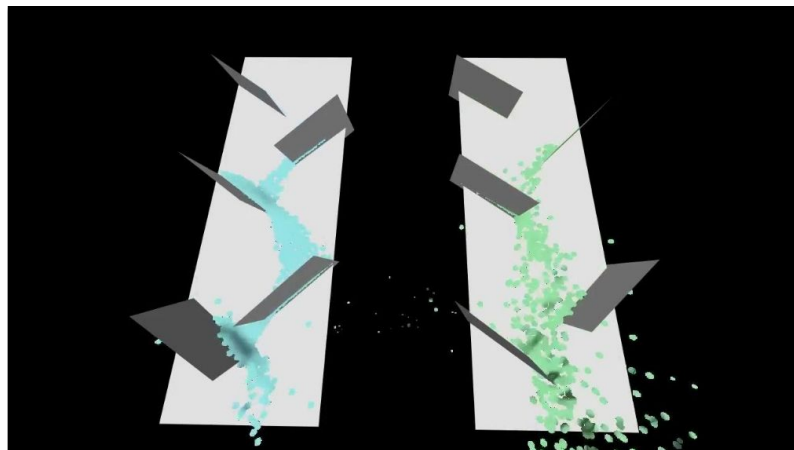
- Computation Irregularity
 - Computing collisions with neighboring particles is a task with a high level of irregularity
- Software/Hardware interfacing
 - Communication needs to happen over a seamless interface to minimize data serialization/unpacking



Computationally Irregular Structure

Testing & Verification

- Quantitative evaluation:
 - Time for each code chunk
 - Time for each frame to render
 - Comparison of resulting data against a golden result
- Qualitative evaluation:
 - Visual inspection to ensure rendered animations still retain “fluid-like” quality



Demo of fluid simulation

Tasks and Division of Labor

Alice	Jeremy	Ziyi
<ul style="list-style-type: none">- Introduce new intuitive UI- Refactor algorithm in a straightforward manner for FPGA translation- Organize data transfer for new UI	<ul style="list-style-type: none">- Manage data transfer between FPGA and CPU- Set up HDMI Interface for Visualization- Optimize fluid simulation algorithm on FPGA- Implement Request scheduler	<ul style="list-style-type: none">- Analyze computational kernel for exploitable parallelism and acceleration opportunities- Set up Vitis HLS workspace and build configurations- Optimize fluid simulation algorithm on FPGA

