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]

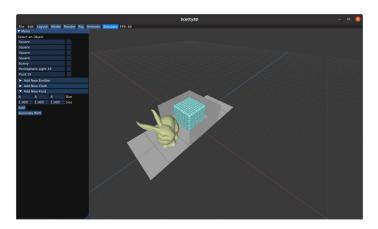


Motivation

 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

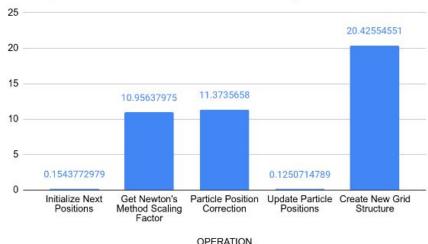


Scotty3D Interface

Performance Requirements

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

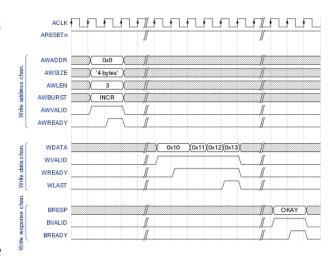
Timing Breakdown of Fluid Simulation Update Iteration



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



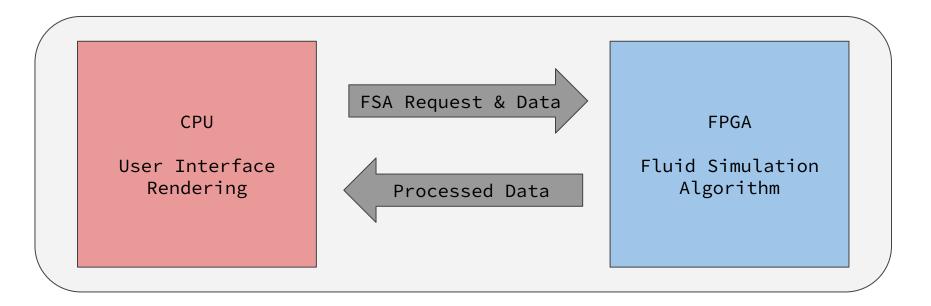
AXI Interface Waveform Diagram

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
 - o Desktops typically run between 50W and 100W

Solution Approach

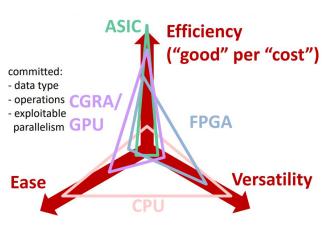
- Accelerate Fluid Simulation Algorithm on the FPGA fabric
- CPU handles the rest of the Scotty3D rendering stack



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

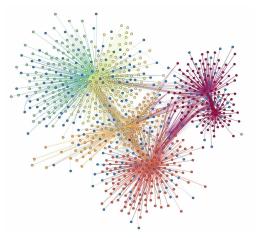
- Exploit thread-level parallelism
 - Unroll loops and instantiate multiple computational units
- Exploit FPGA BRAMS
 - Block RAM high speed versatile memories
 - Large BRAM capacity means we can make multiple copies of data
 - Use this to accelerate neighbor-finding tasks

Algorithm 1 Simulation Loop

```
1: for all particles i do
          apply forces \mathbf{v}_i \leftarrow \mathbf{v}_i + \Delta t \mathbf{f}_{ext}(\mathbf{x}_i)
          predict position \mathbf{x}_{i}^{*} \Leftarrow \mathbf{x}_{i} + \Delta t \mathbf{v}_{i}
 4: end for
 5: for all particles i do
          find neighboring particles N_i(\mathbf{x}_i^*)
 8: while iter < solverIterations do
          for all particles i do
              calculate \lambda_i
          end for
         for all particles i do
              calculate \Delta \mathbf{p}_i
13:
              perform collision detection and response
14:
         end for
         for all particles i do
              update position \mathbf{x}_{i}^{*} \Leftarrow \mathbf{x}_{i}^{*} + \Delta \mathbf{p}_{i}
         end for
19: end while
20: for all particles i do
         update velocity \mathbf{v}_i \Leftarrow \frac{1}{\Delta t} \left( \mathbf{x}_i^* - \mathbf{x}_i \right)
          apply vorticity confinement and XSPH viscosity
          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

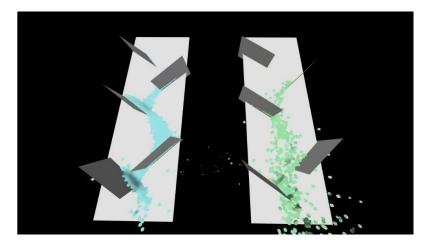


Computationally Irregular Structure

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

- Introduce new intuitive UI
- Refactor algorithm in a straightforward manner for FPGA translation
- Organize data transfer for new UI

Jeremy

- Manage data transfer between FPGA and CPU
- Set up HDMI Interface for Visualization
- Optimize fluid simulation algorithm on FPGA
- Implement Request scheduler

Ziyi

- Analyze computational kernel for exploitable parallelism and acceleration opportunities
- Set up Vitis HLS workspace and build configurations
- Optimize fluid simulation algorithm on FPGA

Schedule

TASK TITLE	TASK OWNER		WEE	К4			WEEK 5				WEEK 6			WEEK 7					WEEK 8					WEEK 9			WEEK 10					WEEK 11					WEEK 12					WEEK 13				WEE		
		М	T W	/ R	F	М	Т	w	R F	М	Т	w	R	М	Т	W	R	F	М	TV	WR	F	М	Т	W	R	F	м	w	R	F	М	T V	V R	F	М	Т	w	R F	М	Т	w	R	FI	VI 1	гν	N F	
Platform																																																
Initial Workflow Setup	All																																															
Figure out memory transfer	Ziyi																																															
Figure out HDMI	Jeremy																																															
Integrate data in / data out	Jeremy																																															
Identify code chunks that complicate FPGA synthesis	All																																															
Redesign C++ code to reduce complexity	Alice																																															
Reimplement C++ code for complexity	Alice/Ziyi																																															
Slack	All	П																																														
FPGA																																																T
Analyze Hardware Resources	Ziyi																																															
Set up Performance Benchmarking	Ziyi																																															T
Get data from interface	Jeremy/Ziyi																																															
Run base kernel on FPGA	Ziyi																																															
Integration	Jeremy/Ziyi																																															
Analyze Performance of first iteration	Ziyi																																															
Slack																																																
Acceleration																																																
Initialize Next Positions	All																																															
Get Newton's Method Scaling Factors	All																																															
Get Particle Position Corrections	All																																															
Update Particle Positions	All																																															T
Reconstruct Grid Structure	All																																															
Slack	All	П																																														
CPU																																																
Set up Performance Benchmarking	Jeremy																																															
Command-Line Interface	Alice																																															T
Data out from interface	Alice/Jeremy																																															
Request Scheduler	Jeremy																																															T
Request Deployment	Jeremy																																									П						
Slack	All																																													İ		
Misc																																																
Final Presentation Slides	All																																															
Slack	All	\Box													\top					\neg	\neg		\top			\neg										Ť				T								