FPGA Based Acceleration of Computational Fluid Flow Simulation on Unstructured Mesh Geometry

Zoltán Nagy, Csaba Nemes, Antal Hiba, András Kiss, Árpád Csík, Péter Szolgay

Computer and Automation Research Institute
Pázmány Péter Catholic University
Széchenyi István University
Introduction

- Challenges in numerical solution of Partial Differential Equations
- High performance Computational Fluid Dynamics (CFD) accelerator
- Arithmetic unit generation and optimization
- Off-chip data access optimization

Results

Conclusions, future work
Acceleration numerical solution of Partial Differential Equations (PDE)

- Wide variety of physical phenomenon
  - sound, heat, elasticity, electrodynamics or fluid flow

- Computational fluid dynamics (CFD)

- Discretization over fine mesh
  - 5M mesh points - Air flow simulation around a car or airplane
  - 200M mesh points – jet engine acoustic modeling

- Weeks of simulation time on clusters
  - Low processor utilization ~10%
  - Weak scalability over ~100 nodes
Workflow

1. Customer (e.g., physicist)
2. Computationally intensive problem defined on a mesh
3. Slow C++ implementation for validation
4. Design strategies
5. VHDL Implementation
   - Memory interface
   - Control
   - Mathematical expression
6. Xilinx Synthesis and P&R tools:
   - Bit code for FPGA and performance evaluation
   - Xilinx constraint file
   - Xilinx IP Core library
Inviscid, Adiabatic, Compressible flows

Euler equations:
\[ \frac{\partial \rho}{\partial t} + \nabla (\rho v) = 0 \]
\[ \frac{\partial (\rho v)}{\partial t} + \nabla \left( \rho vv + I \rho \right) = 0 \]
\[ \frac{\partial E}{\partial t} + \nabla \left( (E + p) v \right) = 0 \]

Total energy is defined as:
\[ E = \frac{p}{\gamma - 1} + \frac{1}{2} \rho v \cdot v \]

Notations:
- \( t \): time
- \( \nabla \): Nabla operator
- \( \rho \): density
- \( v(u, v) \): velocity vector field
- \( p \): pressure
- \( I \): identity matrix
- \( E \): total energy
- \( \gamma \): ratio of specific heats
First order Lax-Friedrichs approximation

\[ U_{i,j}^{n+1} = U_{i,j}^n - \frac{\Delta t}{V_{i,j}} \sum_{f} F_f \cdot n_f \]

\[ F^N = \frac{F_L + F_R}{2} - \left( |\bar{u}| + \bar{c} \right) \frac{U_R - U_L}{2} \]
Data-flow model

1. ...
2. Prod<INSIZE,INSIZE,OUTSIZE> p1,p2;
3. Sum<INSIZE,INSIZE,OUTSIZE> s1;
4. ...
5. p1− >c(signal1);
6. p2− >c(signal2);
7. s1− >a(signal1);
8. s1− >b(signal2);
9. ...

- Mathematical expression implemented in SystemC is converted to a hypergraph.
- Nodes = arithmetic units
- Hyperarcs = connections
Arithmetic unit
Results

Bar chart showing comparison between Global control and Partitioned control. The x-axis represents different control types (LUT, FF, Clock freq.), and the y-axis represents performance ranging from 0% to 300%.

The chart indicates a significant increase in performance when using Partitioned control compared to Global control.
Forward facing step
2D unstructured mesh
Adjacency Matrix: 198,006 nodes

- Bandwidth: 198,006
- Memory:
  - 4 time dependent variables
  - 32 byte/cell, ~6MB
- Node degree: 3
  - 3 x 3byte adjacency list
  - 3 x 2 normal vector coordinate
  - 57 byte/cell, ~10.7MB
  - 3 clk/cell
  - 325MHz clock
  - memory bandwidth 23.5GB/s
  - nonuniform memory access pattern
Renumbering

- Bandwidth: 580
- Memory requirements:
  - 1,160 cell
  - 32 byte/cell: ~36.2kB
- Node degree: 3
  - 3 x 2byte adjacency list
  - 3 x 2 normal vector coordinate
  - 54 byte/cell, ~61.1kB
- 3 clk/cell
- 325MHz clock
- Memory bandwidth 12.7GB/s
Scramjet 3D unstructured mesh
Adjacency Matrix

Scramjet: 210,379 nodes

- Bandwidth: 210,379
- Memory:
  - 5 time dependent variables
  - 40 byte/cell, ~8MB
- Node degree: 4
  - 4 x 3byte adjacency list
  - 4 x 3 normal vector coordinate
  - 108 byte/cell, ~21.6MB
  - 4 clk/cell
  - 325MHz clock
  - Memory bandwidth 28.2GB/s
  - Nonuniform memory access pattern
Renumbering

- Bandwidth: 10,317
- Memory requirements:
  - 20,634 cell
  - 40 byte/cell: ~806kB
- Node degree: 4
  - 4 x 2 byte adjacency list
  - 4 x 3 normal vector coordinate
  - 104 byte/cell, ~2.04MB
  - 4 clk/cell
  - 325MHz clock
- Memory bandwidth: 14.95GB/s
System Architecture

Processor

FIFO

2\textsuperscript{nd} iteration

Processor

FIFO

1\textsuperscript{st} iteration

Processor

FIFO

DMA

Memory Interface and Arbitrer

Off-chip memory

FIFO

DMA

FIFO

DMA

FIFO
Processor Architecture

- Node AddressA
- Write Address
- Node data
- Connectivity descriptor
- Local address generator
- Node AddressB
- AddrA, DOA, DIA
- Memory unit
- AddrB, DOB
- Current node data
- Neighborhood memory unit
- Element descriptor
- Arithmetic unit
- Updated node data
Performance

- Alpha-Data ADM-XRC-6T1
- FPGA: Xilinx XC6VSX475T
  - DSP: 525 (26%)
  - FF: 49,072 (12%)
  - LUT: 34,543 (8%)
  - 3 arithmetic units
- Clock frequency: 325MHz
- 325 million triangle update/s
- 69.22GFLOPs
- 76.3 times speedup

- Intel Xeon E5620 2.4GHz

![Performance Graph](image)
Conclusions, future work

- Supersonic flow simulation
  - High performance FPGA
  - Automatic arithmetic unit generation, partitioning, placement -> high clock frequency
  - Node reordering -> Efficient unstructured mesh handling
  - Nearly two orders (76.3 times) speedup

- Future work
  - Mesh partitioning
  - Multi FPGA
Example: 2D intersection of a Scramjet engine (1.4M grid points)