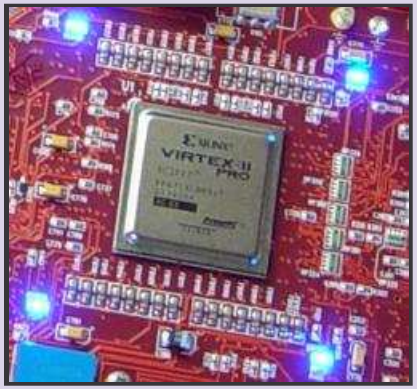
Reconfigurable Computing Aspects of the Cray XD1

Sandia National Laboratories / California

Craig Ulmer

cdulmer@sandia.gov



Cray User Group (CUG 2005)
May 16, 2005

Craig Ulmer, Ryan Hilles, and David Thompson — SNL/CA
Keith Underwood and K. Scott Hemmert — SNL/NM

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Overview

- Cray XD1
  - Dense multiprocessor
  - Commodity parts
  - Good HPC performance
- FPGA Accelerators in XD1
  - Close to host memory
  - *Reconfigurable Computing*
  - Our early experiences

Sandia National Laboratories

Outline

- The Cray XD1 & Reconfigurable Computing
- Four simple applications
  - *DMA, MD5, Sort, and Floating Point*
- Observations
- Conclusions

Sandia National Laboratories

Cray XD1 Architecture

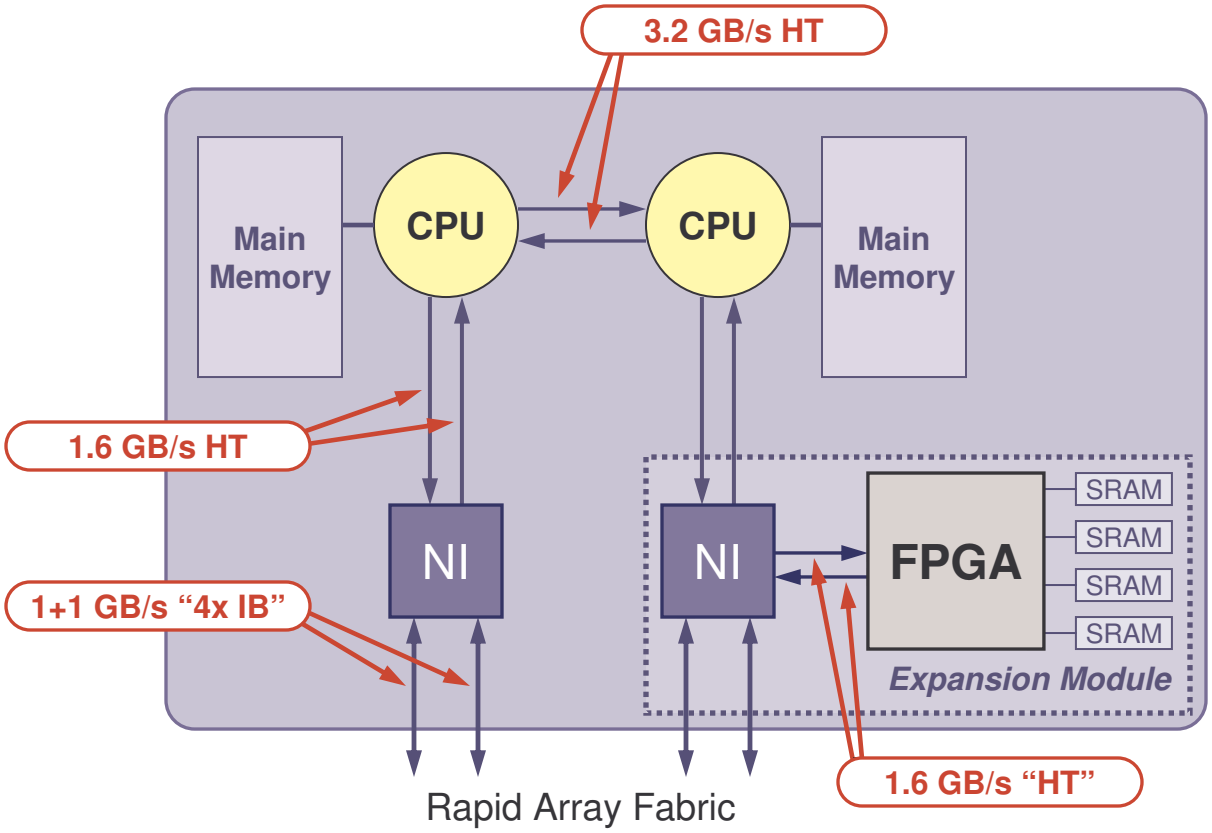
Sandia National Laboratories

Cray XD1

- Dense multiprocessor system
  - 12 AMD Opterons on 6 blades
  - 6 Xilinx Virtex-II/Pro FPGAs
  - InfiniBand-like interconnect
  - 6 SATA hard drives
  - 4 PCI-X slots
  - 3U Rack

Sandia National Laboratories

XD1 Blade Architecture



Sandia National Laboratories

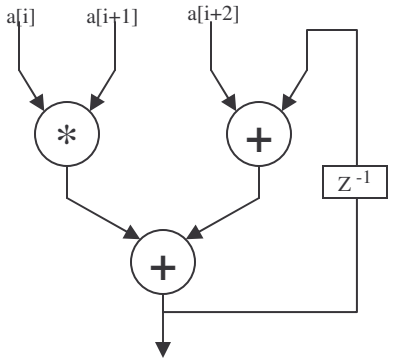
Reconfigurable Computing (RC)

- Use reconfigurable hardware devices to implement key computations in hardware

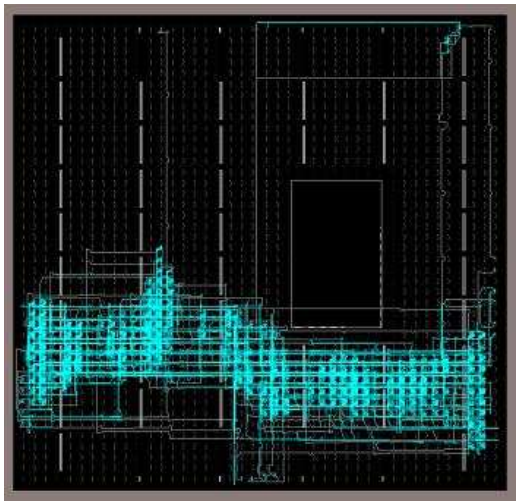
Software

```
double doX( double *a, int n) {
    int i;
    double x;

    x=0;
    for(i=0;i<n;i+=3){
        x+= a[i] * a[i+1] + a[i+2];
        …
    }
    …

    return x;
}
```

Hardware



FPGA



Sandia National Laboratories

Reconfigurable Computing and the XD1

- There have always been three challenges in RC:
  1. System Integration
  2. Capacity
  3. Development Environment

- XD1 is interesting because of (1) and possibly (2)
  - Describe our early experiences
  - Four simple applications to expose performance

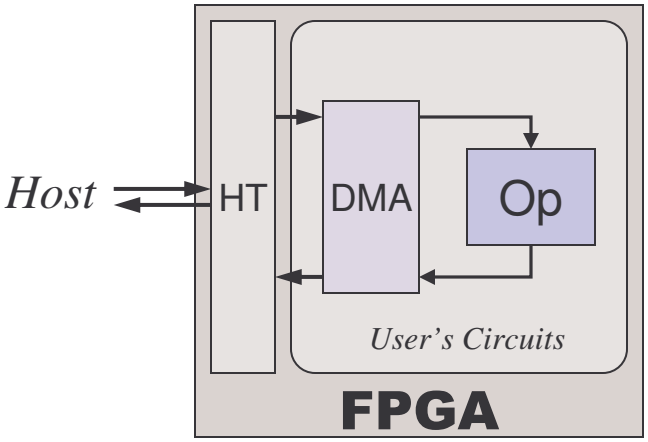
Sandia National Laboratories

Application 1: Data Exchange

Sandia National Laboratories

Data Exchange

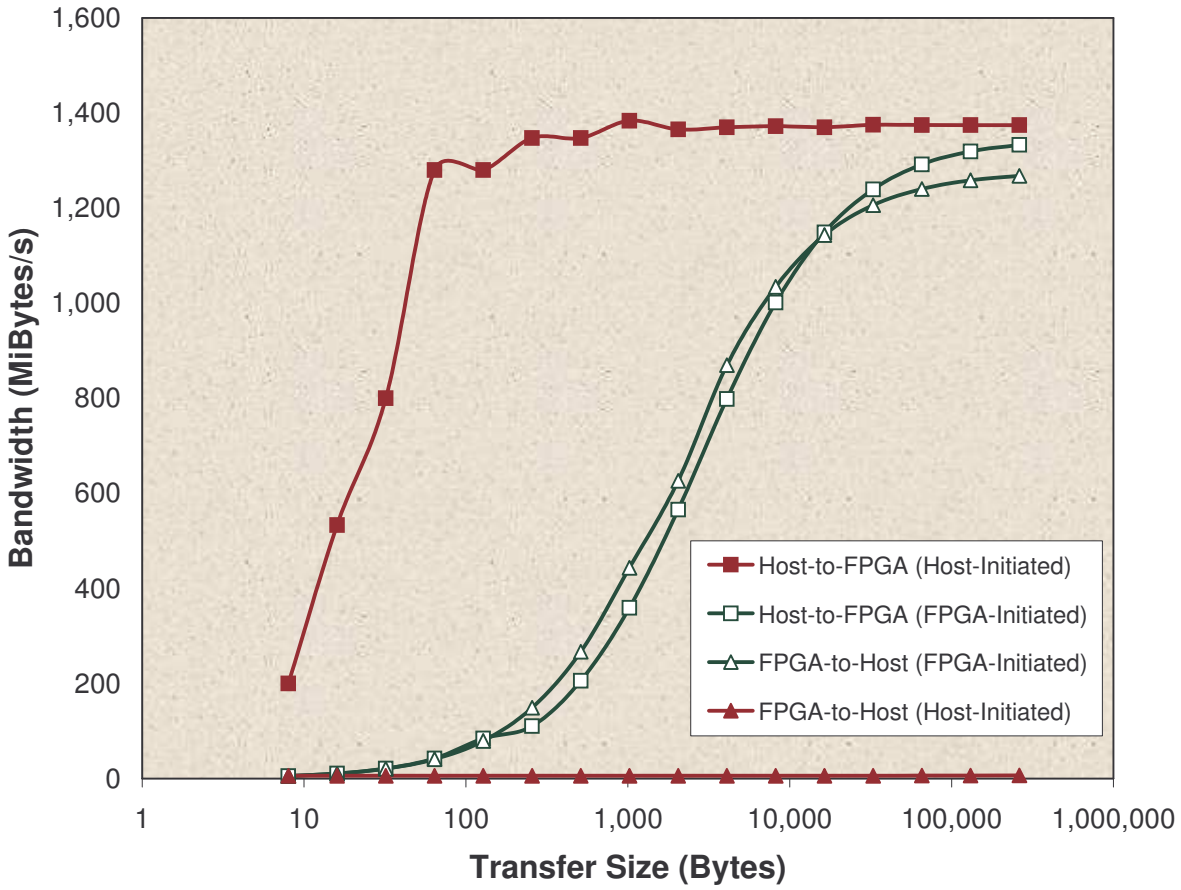
- How fast can we move data into/out of FPGA?
  - Cray provides basic HyperTransport interface
  - Host/FPGA initiated transfers
- HT-Compliance
  - 64B burst boundaries
  - 64b words
- FPGA-Initiated DMA Engine
  - Read/Write large blocks of data
  - Based on memory interface



FPGA-Initiated Transfers

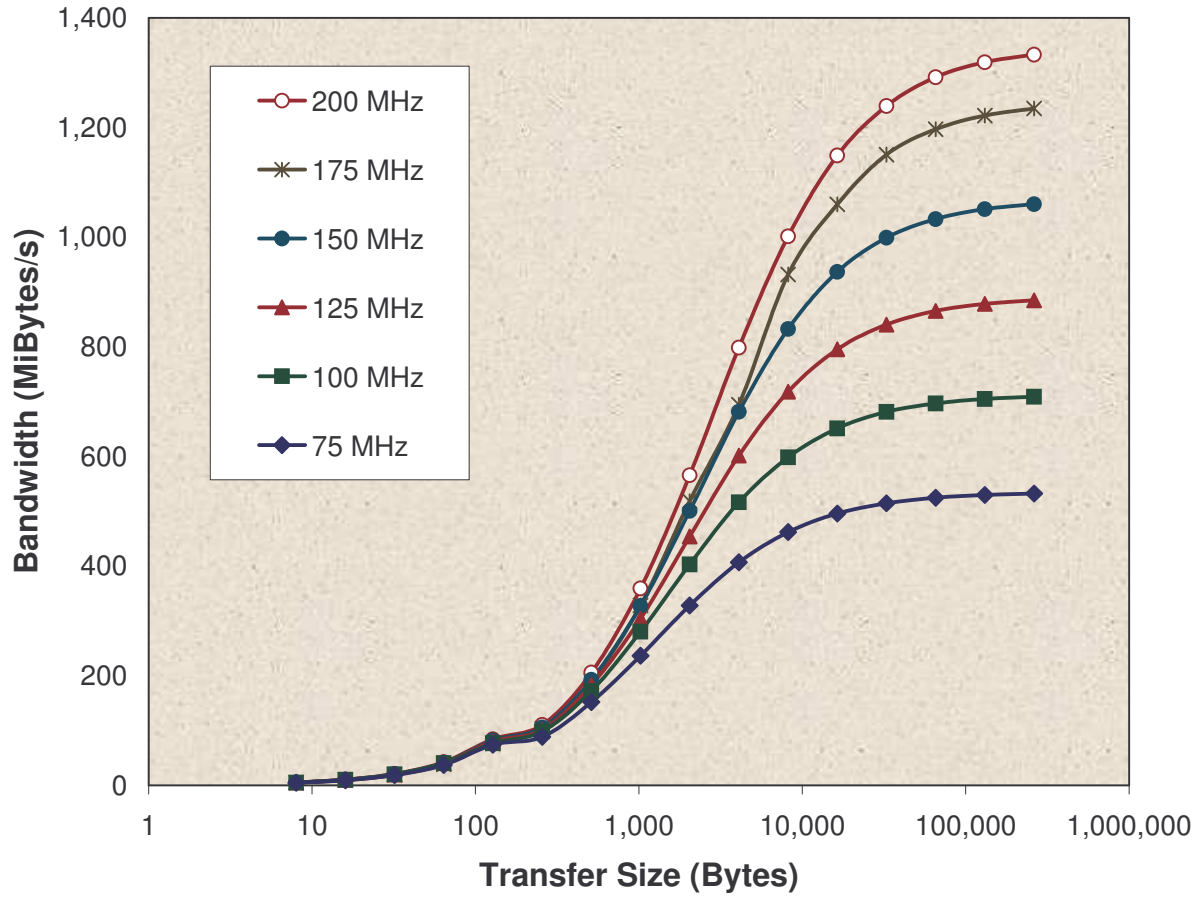
Sandia National Laboratories

Data Transfer Performance



Sandia National Laboratories

FPGA Frequency Affects Performance



Sandia National Laboratories

Application 2: Data Hashing

Sandia National Laboratories

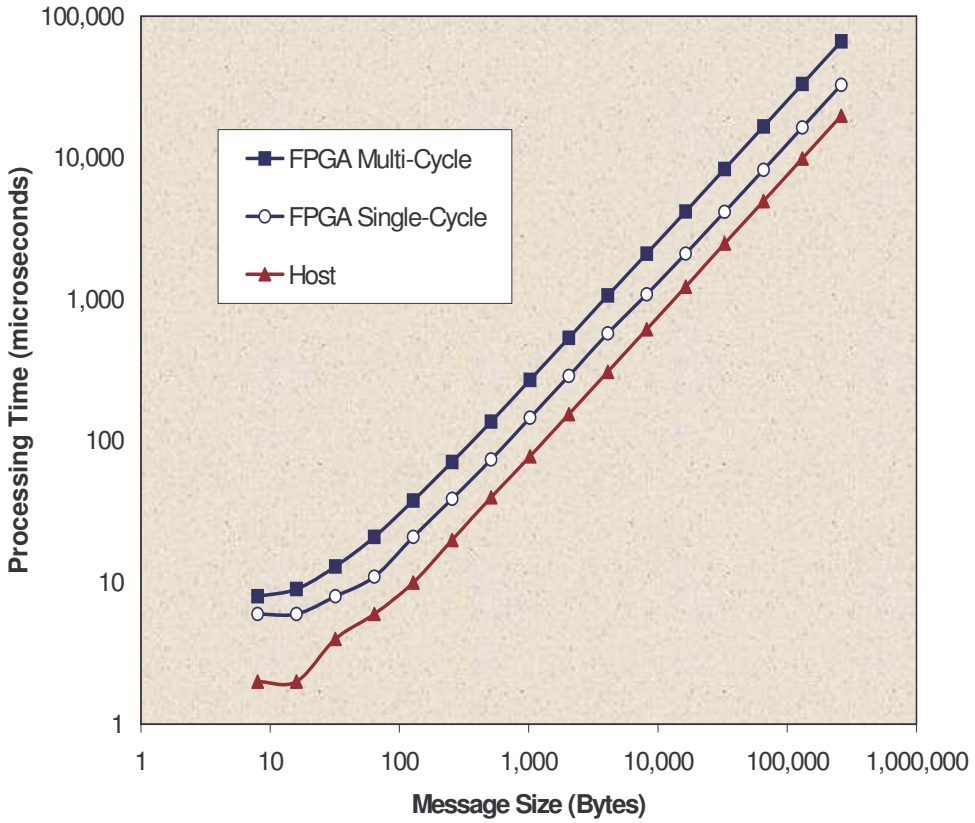
Data Hashing

- Generate fingerprint for block of data
- MD5 Message Digest function
  - 512b blocks of data
  - 64 computations per block
  - Computation: Boolean, add, and rotate
- Two methods
  - Single-cycle computation (64)
  - Multi-cycle computation (320)

Sandia National Laboratories

MD5 Performance

- MD5 is ***serial***
  - Read/Modify fingerprint
  - Host does better
- FPGA clock rates
  - Multi-cycle: 190 MHz
  - Single-cycle: 66 MHz
- Multi-cycle has 5x stages but only 3x clock rate of single-cycle



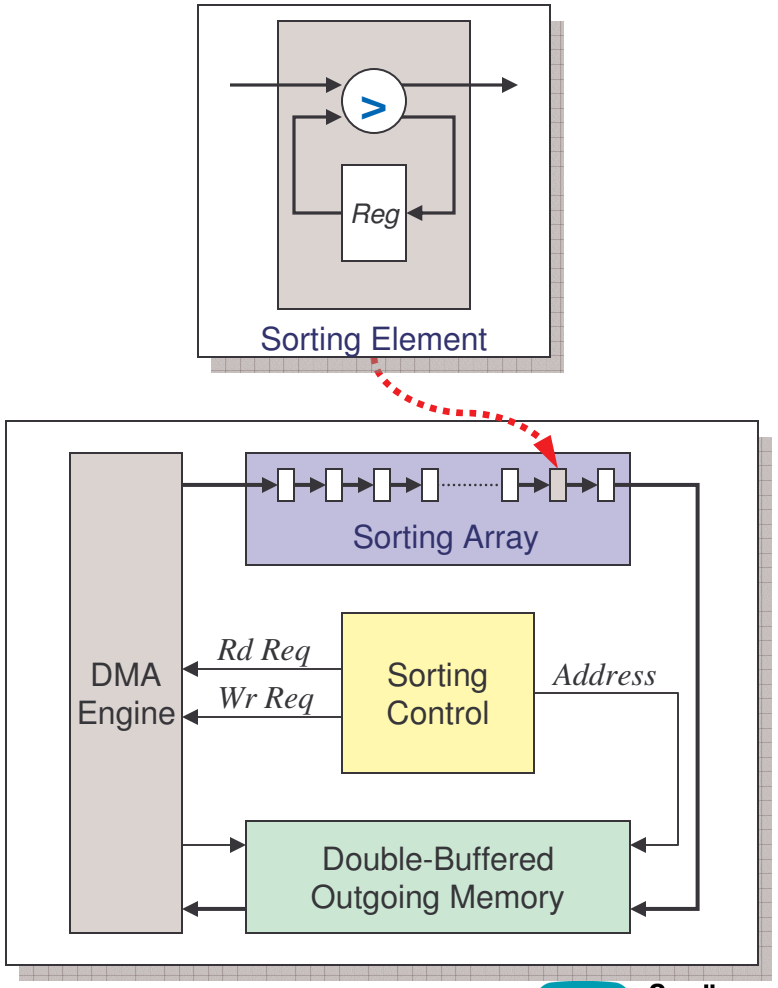
Sandia National Laboratories

Application 3: Sorting

Sandia National Laboratories

Sorting

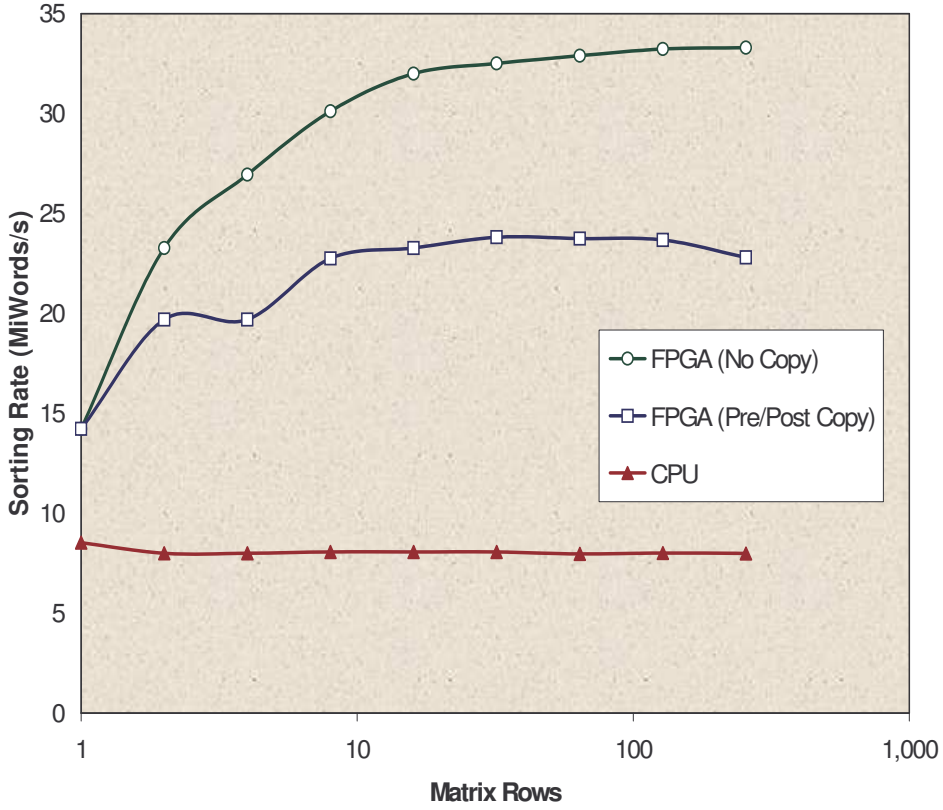
- Sort a matrix of 64b values
- Hardware implementation
  - Sorting element
  - Array of elements
  - Data transfer engine
- Exploits parallelism
  - Do n comparisons each clock
  - Requires 3n clock cycles
- Limited to n values
  - Use as building block



Sandia National Laboratories

Sorting Performance

- V2P50-7 FPGA
  - 128 elements
  - 110 MHz
- Data transfer by FPGA
  - Read/write concurrency
  - Pinned memory
- 2-4x Rate of Quicksort(128)



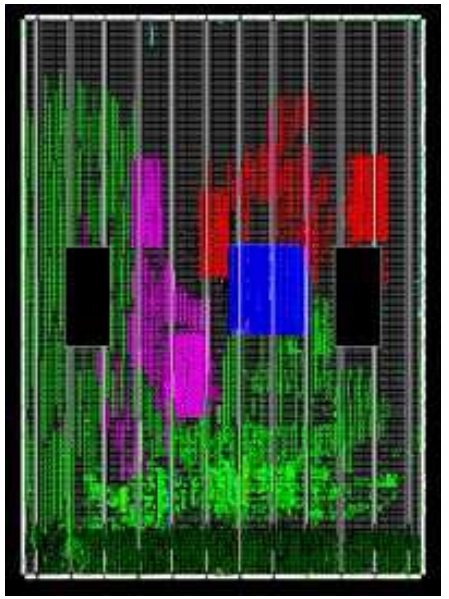
Sandia National Laboratories

Application 4: 64b Floating Point

Sandia National Laboratories

Floating Point Computations

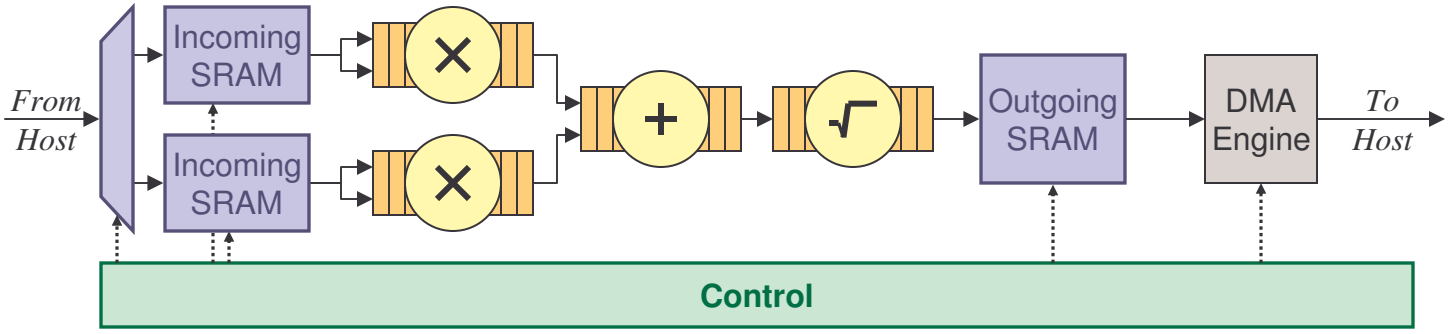
- Floating point has been weakness for FPGAs
  - Complex to implement, consume many gates
- Keith Underwood and K. Scott Hemmert
  - Developed FP library at SNL/NM
- FP Library
  - Highly-optimized and compact
  - Single/Double Precision
  - Add, Multiply, Divide, Square Root
  - V2P50: 10-20 DP cores, 130+ MHz



Sandia National Laboratories

Distance Computation

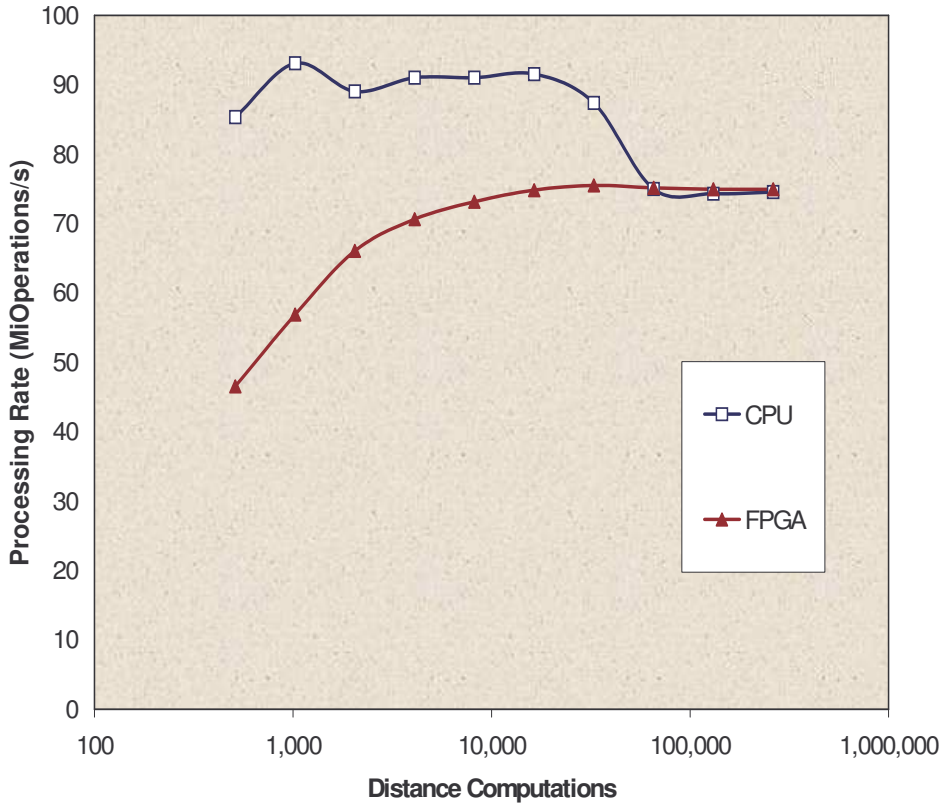
- Compute Pythagorean Theorem: $c=sqrt(a^2 + b^2)$
  - Implemented as simple pipeline (88 stages)
  - Fetch two inputs, store one output each clock
  - Host pushes data in, FPGA pushes data out



Sandia National Laboratories

Distance Performance

- FPGA Design
  - 159 MHz clock rate
  - 39% of V2P50
- CPU and FPGA converge
  - 75 MiOperations/s
  - 1,200 MiB/s of input
- For Square Root and Divide, FPGA can be competitive



Sandia National Laboratories

Observations

- FPGAs have high-bandwidth access to host memory
  - FPGAs can get 1,300 MiB/s – *10x PCI*
  - Still need planning for data transfers
  - Note: host's HT links are 2x
- Some applications work better than other
  - Need to exploit parallelism
- V2P50 is approaching a capacity for interesting work
  - Small number of FP cores
  - Encourage use of larger FPGAs

Sandia National Laboratories

Conclusions

- XD1 is an interesting platform for RC research
  - FPGAs require careful planning
  - High-bandwidth access to memory
  - Sign of interesting HT architectures

- **Acknowledgments**
  - Steve Margerm of Cray Canada for XD1 Support
  - Keith Underwood and K. Scott Hemmert SNL/NM for FP library

Sandia National Laboratories