



USING FIELD PROGRAMMABLE GATE ARRAYS IN A BEOWULF CLUSTER

**Matthew J. Krzych
Naval Undersea Warfare Center**

Problem Description

- ❑ **Building an embedded tera-flop machine**

- ❑ Low Cost
- ❑ Small footprint
- ❑ Low power
- ❑ High performance

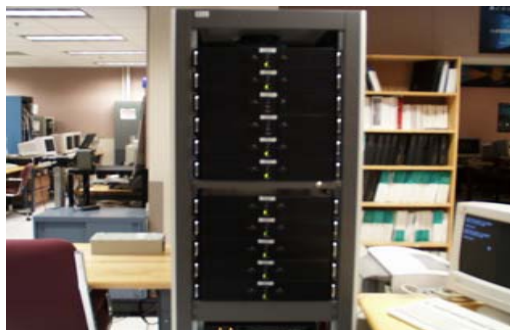
- ❑ **Utilize commercially available hardware & software**

- ❑ **Application:
Beamform a volume of the ocean**

- ❑ Increase the number of beams
from 100 to 10,000,000



On February 9, 2000 IBM formally dedicated [Blue Horizon](#), the teraflops computer. [Blue Horizon](#) has 42 towers holding 1,152 compute processors, and occupying about 1,500 square feet. Blue Horizon entered full production on April 1, 2000.



**Beowulf
Cluster**

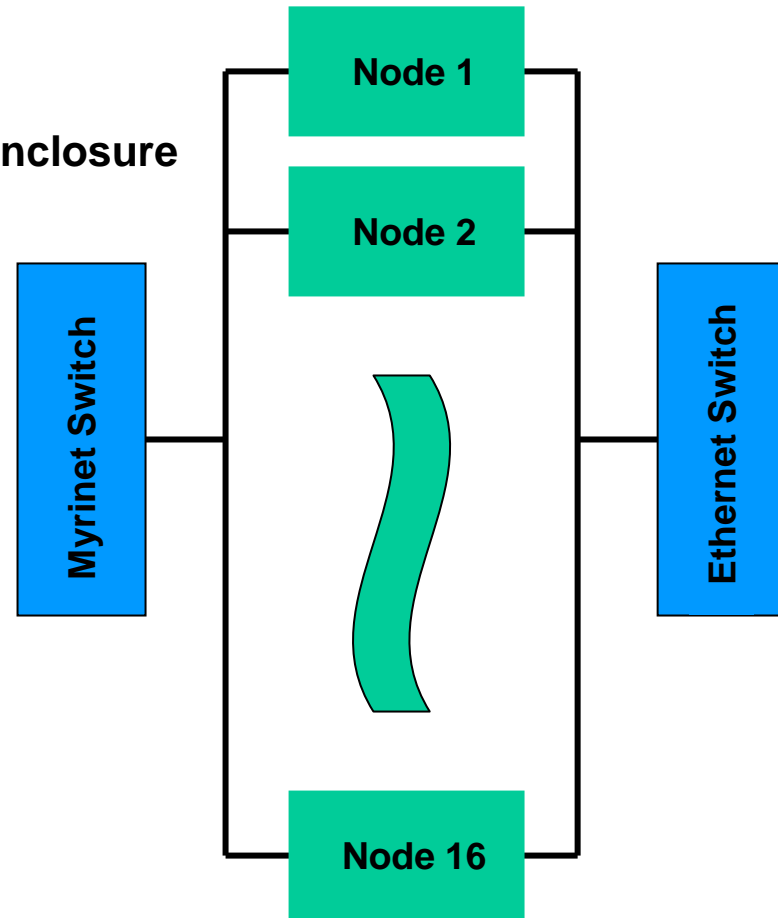
System Hardware

- ❑ **16 Node Cluster**
 - ❑ AMD 1.6 GHz and Intel Pentium 2.2 GHz
 - ❑ 1 to 4 GBytes memory per node
 - ❑ 2U & 4U Enclosures w/ 1 processor per enclosure
 - ❑ \$2,500 per enclosure ¹.

- ❑ **8 Embedded Osiris FPGA Boards**
 - ❑ Xilinx XC2V6000
 - ❑ \$15,000 per board ¹.

- ❑ **Myrinet High Speed Interconnect**
 - ❑ Data transfer: ~250 MBytes/sec
 - ❑ Supports MPI
 - ❑ \$1,200 per node ¹.
 - ❑ \$10,500 per switch ¹.

- ❑ **100 BASE-T Ethernet**
 - ❑ System control
 - ❑ File sharing



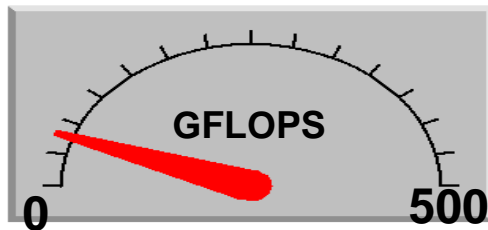
➡ **Total Hardware Cost¹: \$190K** ⬅

¹. Cost based on 2001 dollars. Moore's Law asserts processor speed doubles every 18 months. 2004 dollars will provide more computation or equivalent computation for fewer dollars.

Lessons Learned

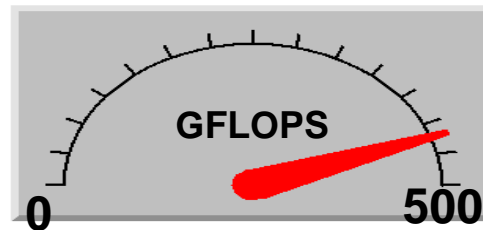
❑ WITHOUT hardware accelerator

- ❑ 16 nodes (2.2 GHz)
- ❑ 5 GFLOPS sustained
 - Single precision



❑ WITH hardware accelerator

- ❑ 8 FPGA boards
- ❑ 500 GFLOPS
 - Fixed point
 - Pipelining
 - Parallelism



❑ Beowulf Cluster

- ❑ Flexibility / robustness
 - Supports heterogeneous hardware
 - Run-time selection of processors, functions, & system parameters
- ❑ Scalability
 - Add / remove hardware assets
 - Add / remove functionality

❑ MPI

- ❑ Facilitates flexibility & scalability
- ❑ Runs on multiple hardware platforms & operating systems
- ❑ Supports multiple communication schemes (point-to-point, broadcast, etc.)