



Big Data Meets High-Performance Reconfigurable Computing

UF Workshop on Dense, Intense, and Complex Data



Alan George
CHREC Center Director

Herman Lam
CHREC Center Associate Director

UF UNIVERSITY of FLORIDA



THE GEORGE WASHINGTON UNIVERSITY WASHINGTON DC

Virginia Tech

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

BYU

BRIGHAM YOUNG UNIVERSITY

What is CHREC?

NSF Center for High-Performance Reconfigurable Computing

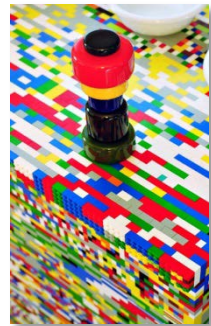
- Industry/University Cooperative Research Center (I/UCRC)
 - Unique US national research center, operational since January 2007
- CHREC is both national research center and consortium
 - **Research base:** Leading ECE/CS research groups @ four major universities
 - University of Florida (lead)
 - George Washington University
 - Brigham Young University
 - Virginia Tech
 - **Members:** Industry & government organizations
 - Often cited by NSF as one of top national centers
 - e.g., CHREC won 2012 Schwarzkopf Prize for Technology Innovation

1. AFRL Munitions Directorate
 2. AFRL Sensors Directorate
 3. AFRL Space Vehicles Directorate
 4. Altera
 5. AMD
 6. Arctic Region Supercomputing Center
 7. Convey Computer
 8. Data IO
 9. Draper Lab
 10. GiDEL
 11. Harris
 12. Honeywell
 13. IBM
 14. Lockheed Martin MFC
 15. Lockheed Martin SSC
 16. Los Alamos National Laboratory
 17. MIT Lincoln Laboratory
 18. NASA Goddard Space Flight Center
 19. NASA Kennedy Space Center
 20. NASA Langley Research Center
 21. National Instruments
 22. National Security Agency
 23. Office of Naval Research
 24. Sandia National Laboratories
 25. SEAKR Engineering
 26. Space Micro
 27. Texas Instruments
 28. Xilinx
- CHREC Members**
Industry & government organizations

Conventional vs. Reconfigurable Computing

■ Conventional computing

- Each *app must conform to predefined, fixed hardware of target*
 - e.g. CPU, DSP, or GPU; fixed cores, memory, interconnect, & I/O structures
- GOOD performance when app maps well to predefined hardware
 - But poor performance when app maps poorly
- Programming simplicity with single core.
 - But programming complexity in scaling beyond one core



■ Reconfigurable computing

- Computing with hardware-reconfigurable circuits, devices, systems
 - Architecture can be *adapted to match unique needs* of each app or task
- GREAT performance when app maps poorly to predefined H/W
 - **Customized** parallelism, data precision, operations, memory structure, interconnect
- Much more *energy-efficient* than fixed-logic processors

Big-Data in Key Science Domains

Bioinformatics

- **Sequencer output increases by 10x every 2 years!**
 - Moore's Law: CPU performance doubles every 2 years
 - Kryder's Law: storage quadruples every 2 years



Computational finance

- **Financial markets are significant producers of Big Data**
 - Trades, quotes, earning statements, statistical releases, polls, etc.
- **Our current emphasis is on risk analysis and management**
 - Improve enterprise transparency, auditability, and oversight of risks
 - New regulatory and compliance requirements on risk reporting



Signal, image, & video processing

- **Escalation in sensor fidelity and diversity**
 - Explosion in data: limited time (real-time), power, space, & cost
- **Vital in areas such as:**
 - Computer/machine vision, medical imaging/diagnosis, face recognition, autonomous navigation, tracking, satellite imaging, etc.



Novo-G Supercomputer

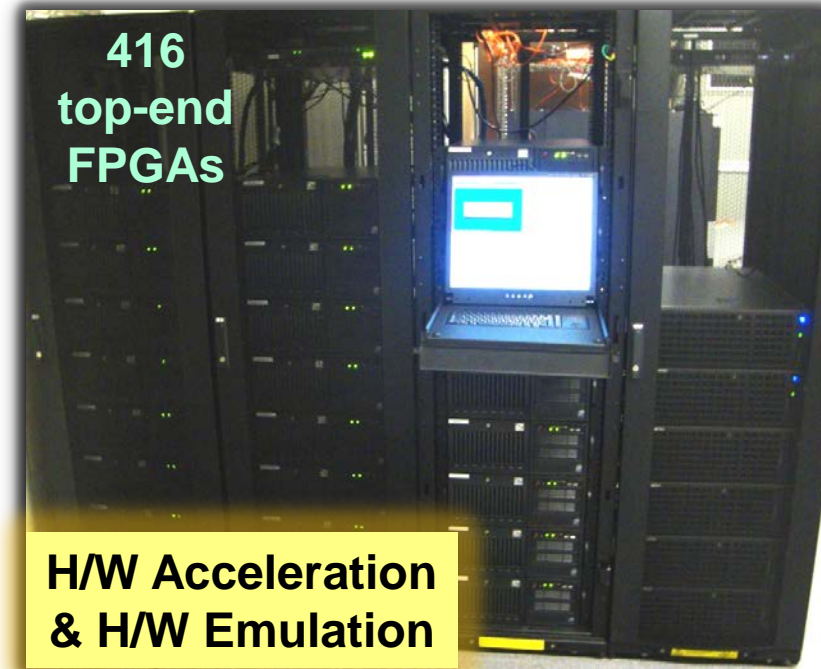


- Developed and deployed at CHREC
 - Hardware acceleration of parallel applications
 - Hardware emulation of next-gen systems & apps
- Most powerful RC machine in known world
 - For some apps & uses, may be fastest computer of *any kind* in world!
 - O(1000) of times less cost, size, power, & cooling than massive supercomputers
- Award-winning system
 - 2012 Schwarzkopf Award
 - 2010 HPCwire Award
- New system expansion
 - Addition of top-end 28nm FPGAs (32)
 - Stratix-V D8 devices in PROCe-V boards
 - Focus on high-performance interconnect
 - 3D-torus & 5D-hypercube FPGA topologies



Novo-G Annual Growth

- 2009: 96 top-end Stratix-III FPGAs, each with 4.25GB SDRAM
- 2010: 96 more Stratix-III FPGAs, each with 4.25GB SDRAM
- 2011: 96 top-end Stratix-IV FPGAs, each with 8.50GB SDRAM
- 2012: 96 more Stratix-IV FPGAs, each with 8.50GB SDRAM
- 2013: 32 top-end Stratix-V FPGAs (4x4x2 torus or 5D h-cube)



BioRC Research Highlights

- **Smith-Waterman on Novo-G**
 - Algorithm for optimal **local alignment** of DNA or RNA sequence pairs
 - Current configuration: **650 PEs per FPGA** @ 125MHz (81 GCUPS limit)
- **CHREC BLAST Toolset on Novo-G**
 - **Novo-BLAST:**
 - **High performance:** hardware-accelerated BLAST algorithm
 - **Output identical to NCBI-BLAST**
 - **BSW (BLAST-Wrapped Smith-Waterman):**
 - **Max. sensitivity, with runtimes on order of NCBI-BLAST**

Baseline: Human X Chromosome v 19200, length 650 Seqs
Software Runtime: 5,481 CPU hours on 2.4GHz Opteron

# FPGAs	Runtime (sec)	Speedup *
1	23,846	827
4	5,966	3,307
96	250	78,926
128	188	104,955
192 (est.)	127	155,366

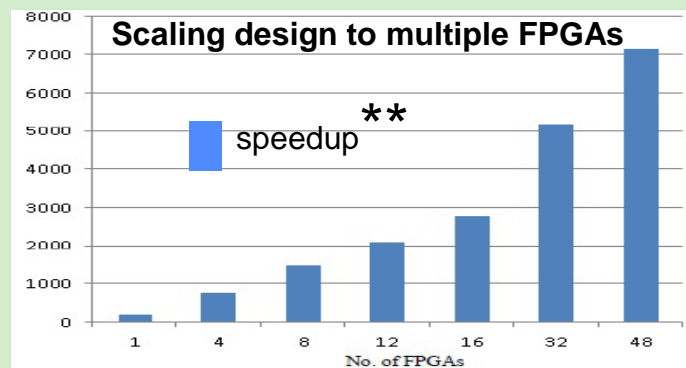
* *Stratix-III E260 FPGAs at 125 MHz*

Performance equivalent to 155,366 Opteron cores!



FinRC Research Highlights

- **FinRC app for derivative pricing**
 - Simulation model for **multi-asset barrier** options using **Heston** dynamics
- **FinRC OpenCL studies**
 - **Standard benchmark** defined by Securities Technology Analysis Center (STAC)
 - Productivity studies using Altera **OpenCL**



** *Stratix IV E530 FPGA at 125 MHz; C-based SSE2 optimized baseline; On 1 core of Intel Xeon E5-2687 @ 3.1 GHz; Maturity = 10 years; 1,000,000 paths; 16 asset option*

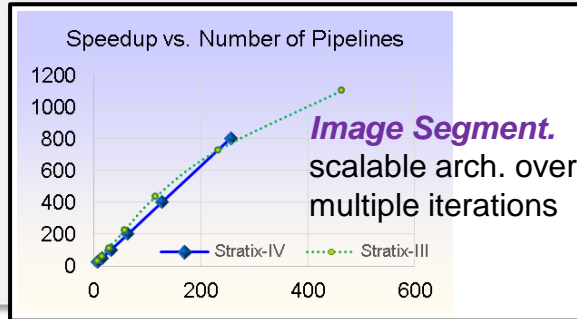
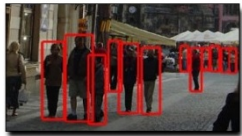
DspRC Research Highlights

Image Segmentation (Information Extraction)

- **Goal:** Locate & separate objects from background

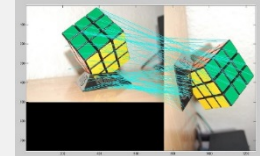


Object Recognition



Speeded Up Robust Features = SURF (Decision Making)

- **Goal:** Extract features and recognize objects/surroundings



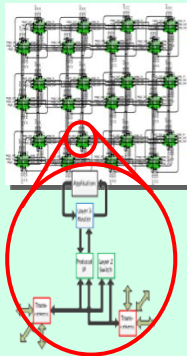
Motivation for:
Unsupervised + real-time image and video processing for autonomous applications

Autonomous Navigation

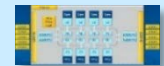


Architectural Studies for Big Data

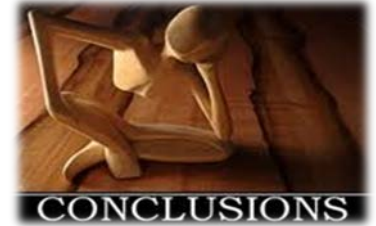
- Explore **adaptive networks** to accelerate communication-intensive Big-Data apps
 - Efficient FPGA-to-FPGA multi-dimensional, backend networks
 - e.g., *Adaptable* 3D-torus or 5D-hypercube



- **Many-core** and **hybrid-core** exploration for Big-Data
 - Tileria (up to 72 cores)
 - Intel Xeon Phi (60 cores)
 - TI KeyStone hybrid-core (up to 4 ARM-A15 cores; 8 DSPs)
 - Xilinx ZYNQ hybrid-core (Dual ARM-A9 cores; FPGA)



Conclusions



- Big-Data challenges and big opportunities
- Reconfigurable computing (RC) offers unique solution to key Big-Data problems
- CHREC, as a national research center & consortium, at forefront of HPRC
 - Novo-G reconfigurable supercomputer
 - Focusing on Big-Data bottlenecks in key science domains: BioRC, FinRC, DspRC
 - Architectural studies: many-cores, hybrid-cores