ORNL Field-Programmable Gate Array (FPGA) Research Speeds HPC up to 100X

Presented by

Olaf O. Storaasli

Future Technologies Group
Computer Science and Mathematics Division

Explore FPGAs for future ORNL HPC

Industry view

More petaflops at reduced power

"*After exhaustive analysis, Cray concluded… hardware accelerators (e.g., FPGAs* or *ClearSpeed co-processors) create the* ***greatest opportunity*** *for application acceleration*."

THE SUPERCOMPUTER COMPANY

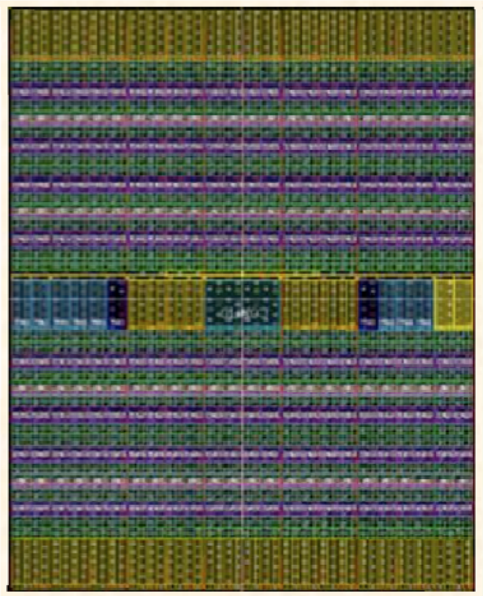
Steve Scott, CTO HPCWire

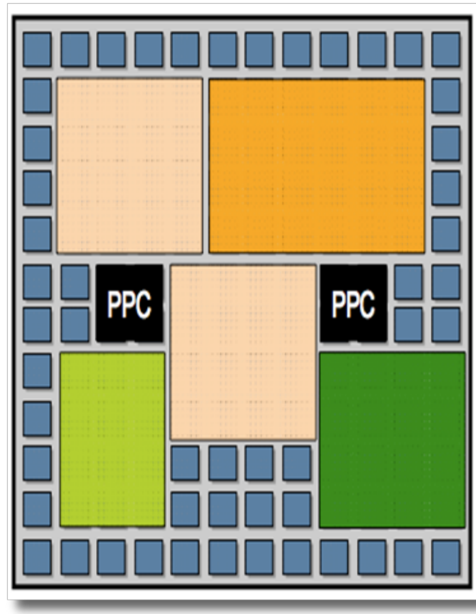
Contents

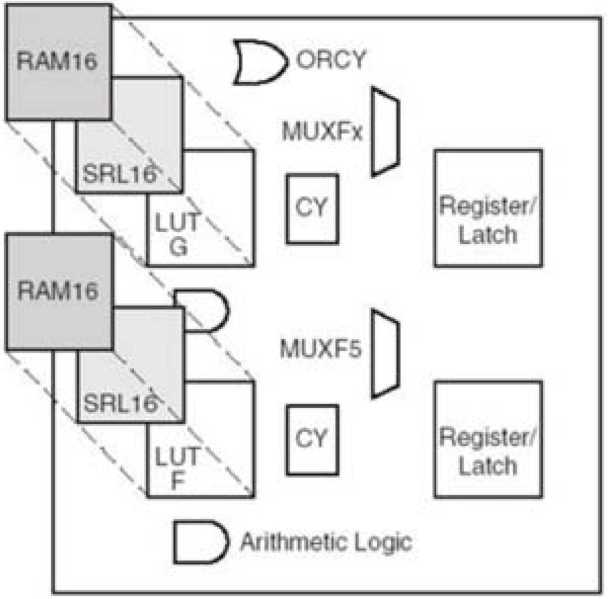
- **Background: Why FPGAs?**
- **ORNL success: FPGA systems, tools and up to 100X speedup**
- **Partners: XILINX Research Lab SRC mitrion**

FPGAs: Your “custom chip”

Future Technologies Group







Xilinx Virtex4 FPGA: 25K slices

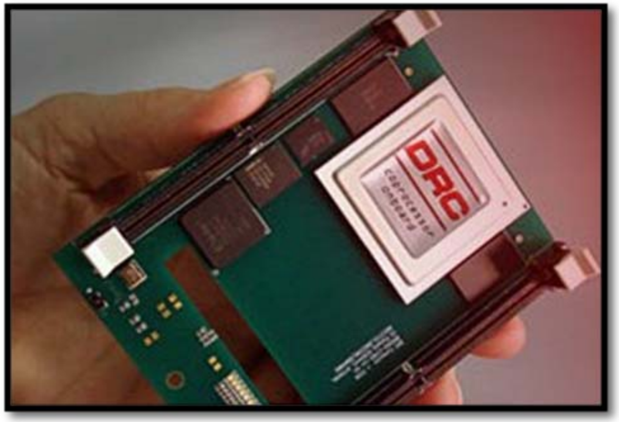
- **Tailor logic array to your application.**
- **On-chip RAM, multipliers and PowerPCs.**
- **FastIO: Gigabit transceivers/DSP blocks.**
- **100–1000 operations/clock cycle.**

FPGA Logic slice
(MiniCPUs)

Why FPGA accelerators?

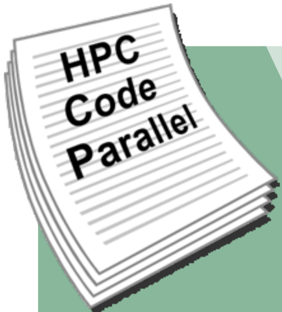
- **Performance—optimal silicon use:**
  maximize parallel ops/cycle.
- **Rapid growth—cells, speed, I/O.**
- **Power—1/10th CPUs.**
- **Flexible—tailor to application.**

Cray FPGA accelerators



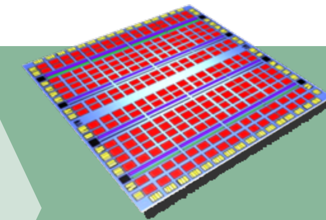
HPC code (STSWM) port to FPGAs

ORNL-Xilinx Collaboration



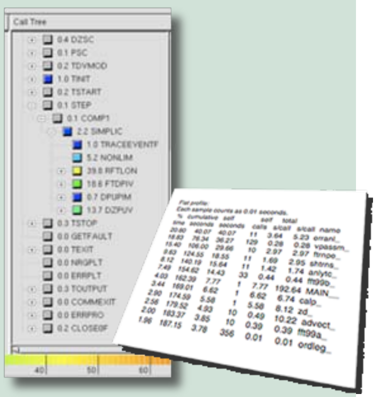
HLL developer profiles

HLL compiler
CHiMPS, Mitrion
(FPGA Tools Inside)

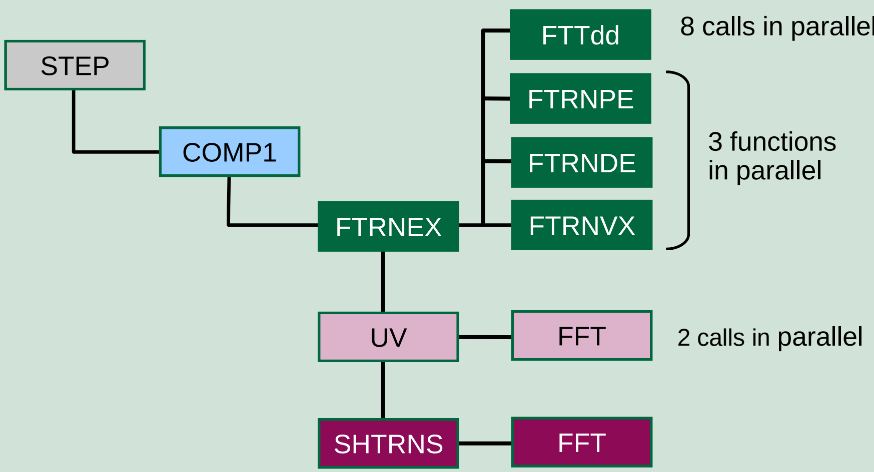


FPGA speedup

Profile

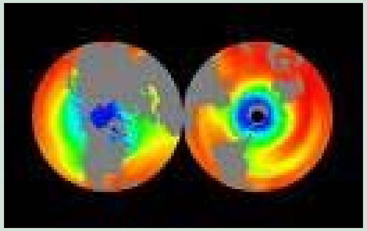


Find parallelism: 80% FFTs



Goal

More GF/$ GF/Watt

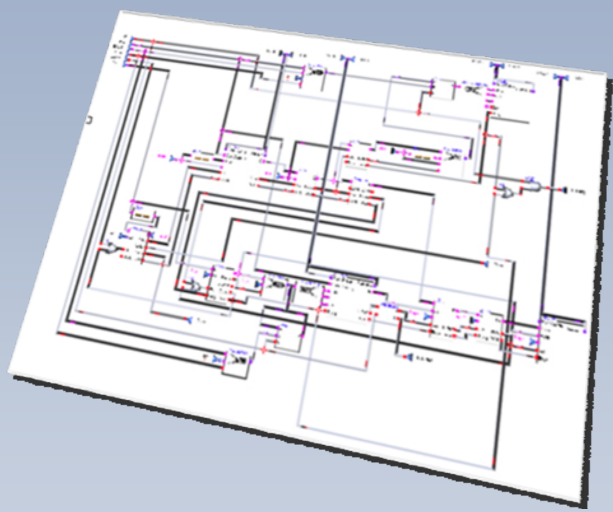


Model faster

FPGA coding options

Future Technologies Group

Gauss matrix solver



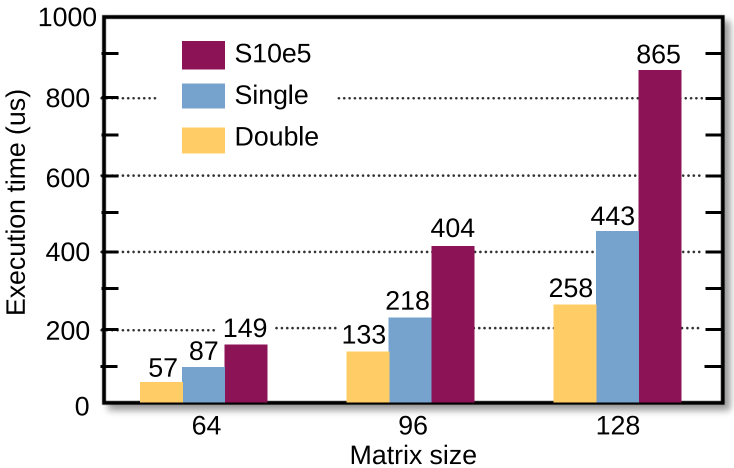
Viva: Graphical icons—3-dimensional

Compiler, simulator, and debugger

MitrionC: Text/flow—1-dimensional

+ Carte/SRC, CHiMPS-VHDL/Xilinx, DSPlogic

37X* LU decomposition speedup
10X for matrix equation solver



Design	Double FP	Single FP	S10e5
PE amount	8	16	32
Max size	128	256	256
Achievable frequency	120 MHz	150 MHz	150 MHz
Slices	27,005 (57%)	14,792 (59%)	14,730 (62%)
BRAMs	68 (29%)	129 (55%)	65 (28%)
MULT18X18	128 (55%)	64 (27%)	32 (13%)

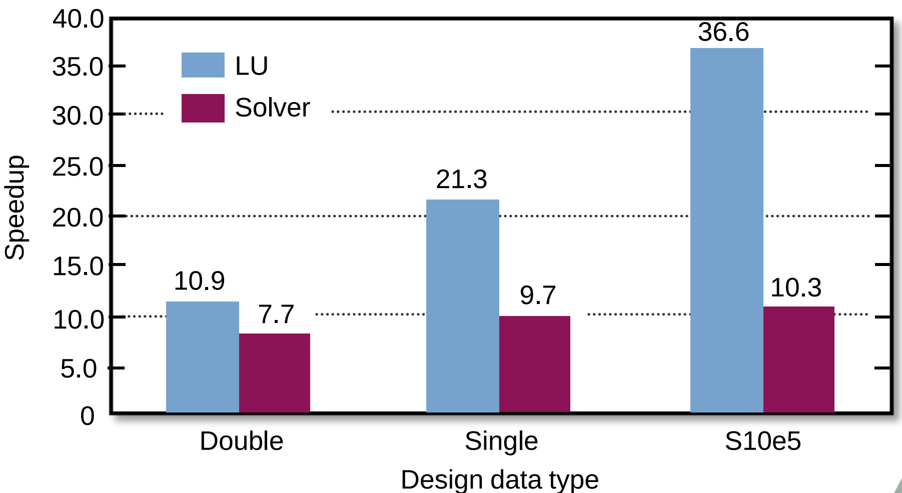
Benefits:

High performance of LP arithmetic.

High-precision accuracy.

Speedup increases with matrix size

as LU dominates calculations.

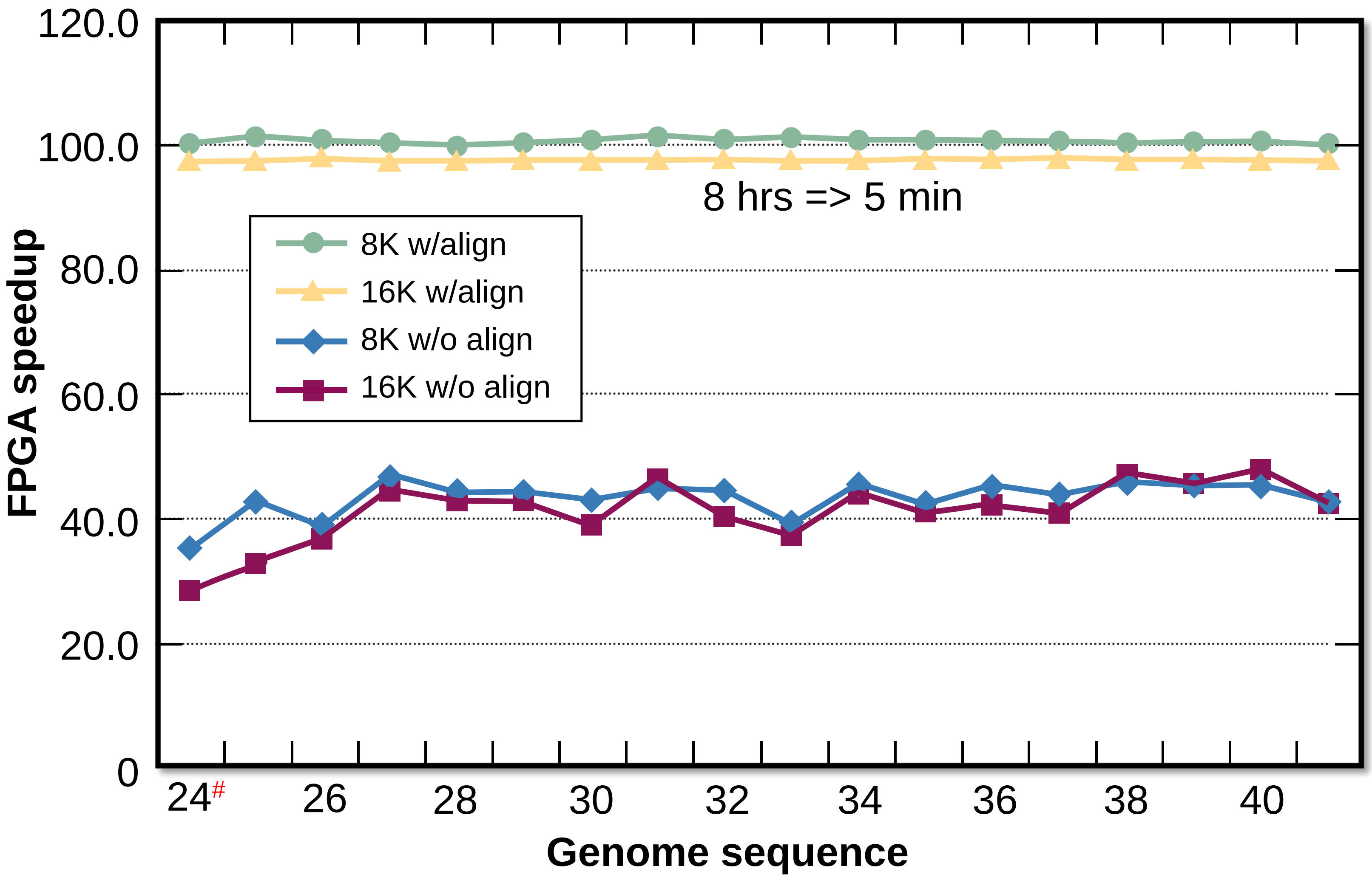


First mixed-precision LU and solver for FPGAs

*2.2 GHz Opteron

100X* speedup

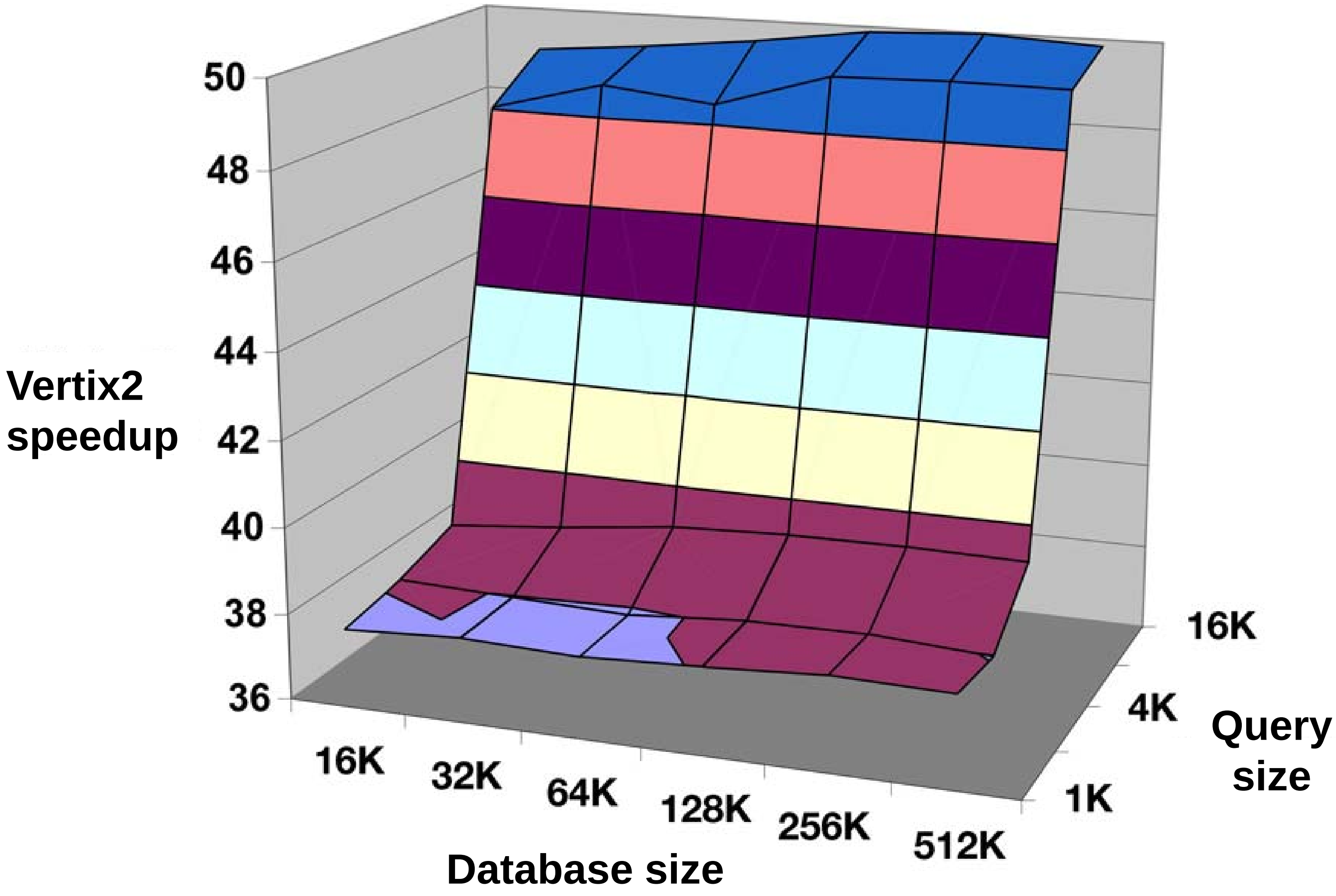
Bacillus anthracis human DNA sequencing



*Virtex-4 FPGA vs 2.2 GHz Opteron on Cray XD1

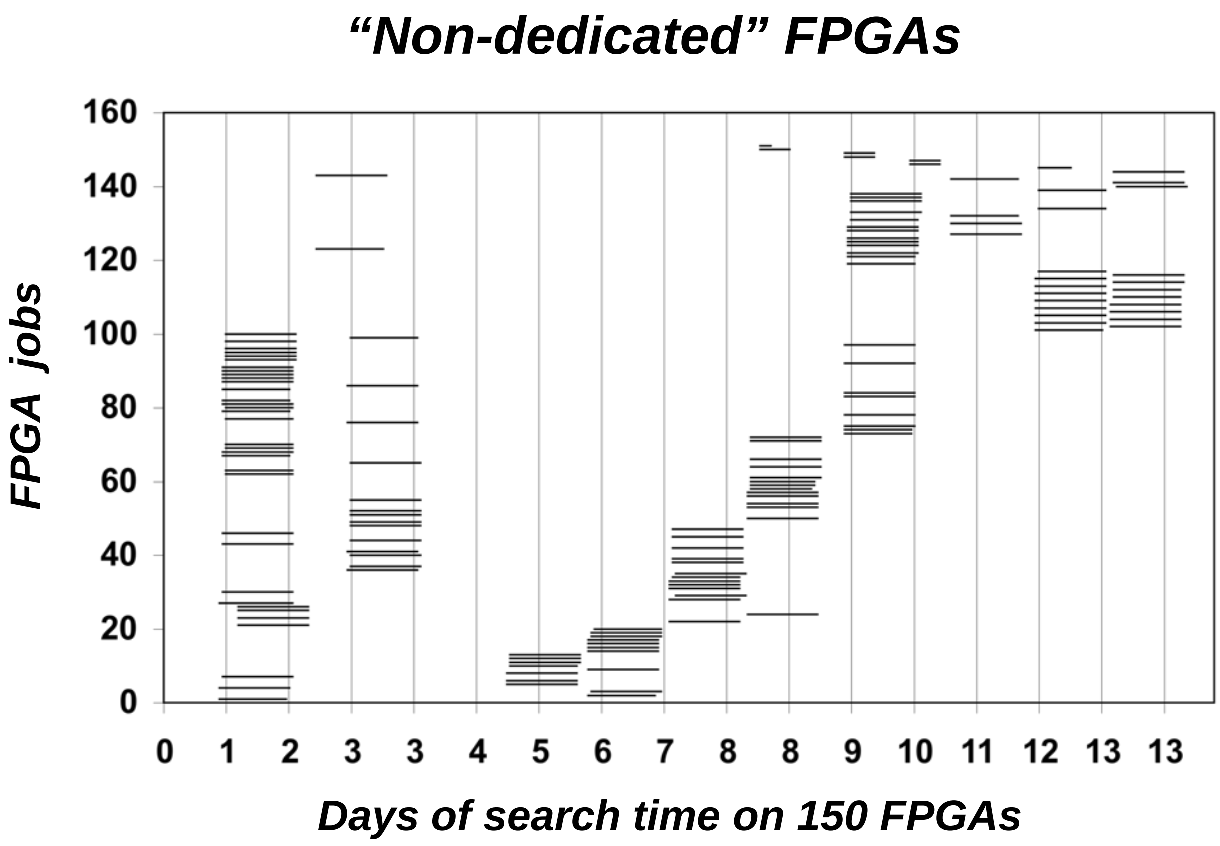
24= Sequence AE17024

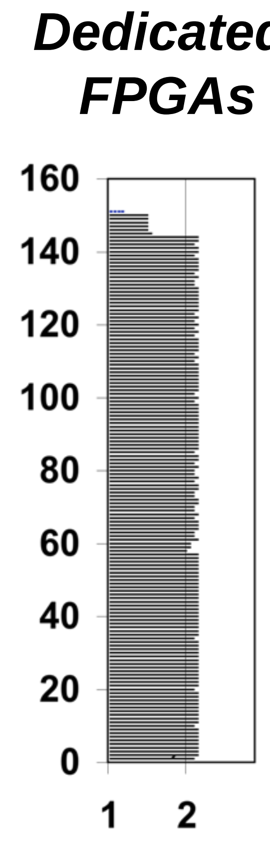
FPGA speedup with query size



DNA sequencing* time on 150 FPGAs

*Human-mouse DNA compare (FASTA**)**





DNA Sequence Speed* on 150 FPGAs

** State-of-the-art:* **Giga Cell Updates Per Second (GCUPS)**

- **DNA characters: Human = 155 million, mouse = 165 million.**

Total compares = 155M × 165M × 1062 × 2 = 51x10^{15} **cell updates.**

- **Sequential FPGAs take 11,923,200 s** (138 days)
  ==> 51 × 10^{15}/11,923,200 **= 4.3 TCUPS (*Tera CUPS*)**
- **Parallel (actual) = 1,114,560 s** (12.9 days) **==> 46 TCUPS.**
- **Parallel (dedicated) = 86,400 s** (1 day) **==> 605 TCUPS.**

Summary

Speedup* on 1 FPGA:

10X for general matrix equation solution.
100X *for DNA sequencing.*

Speedup* on 150 FPGAs *for DNA Sequencing:*

1 Opteron ==> 18 years   150 Opterons ==> 6 weeks
1 FPGAv2 ==> 5 months   150 FPGAs ==> 1 day ***49X speedup***

==> 7,350X speedup over one Opteron (VirtexIIs)
==> 14,700X speedup (Virtex4s)

*Compared with one 2.2 GHz Opteron

Contact

Olaf Storaasli

Future Technologies Group
Computer Science and Mathematics Division
Olaf@ornl.gov
Google Olaf ORNL

Acknowledgment:

Thanks are extended to the Naval Research Lab
for use of its Cray XD1 with 150 FPGAs