



# **Beyond 100x Speedup with FPGAs**

## **Cray XD1 I/O Analysis**

**Dr. Olaf O. Storaasli**  
**Future Technologies Group**  
**Computer Science & Mathematics Division**  
**Oak Ridge National Laboratory**  
**&**  
**Dave Strenski, Cray Inc.**

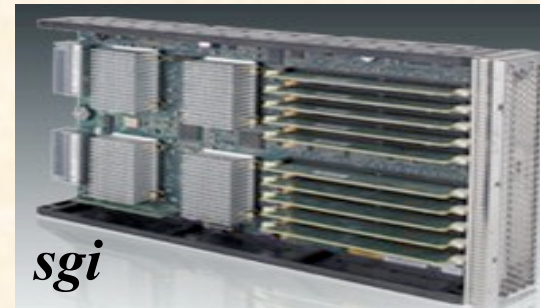
***Cray User Group, Atlanta 5-5-09***

# 3 FPGA Generations: Moving toward HPC

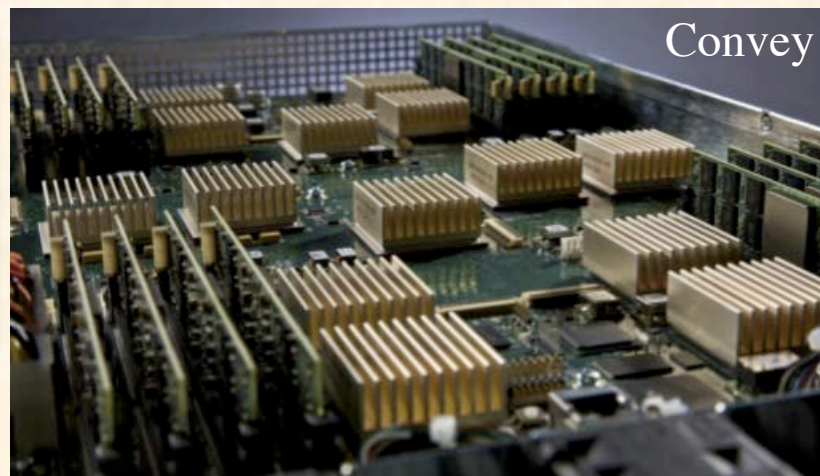
**PCI:** ANS, DSP => HPEC



**HT:** Cray XD1, *sgi*, SRC, ...



**Socket:** Cray XT5h (DRC, XtremeData), Convey



# ***ORNL Cray XD1 with Xilinx Virtex2 FPGAs***



## Why FPGAs?

- **Performance:** optimal silicon use, *maximize parallel ops/cycle*
- **Rapid growth:** Cells, Speed, I/O
- **Power:** 1/10th CPUs
- **Flexible:** *tailor* to application
- **Advances:** Telecom industry spinoff

## Why not FPGAs?

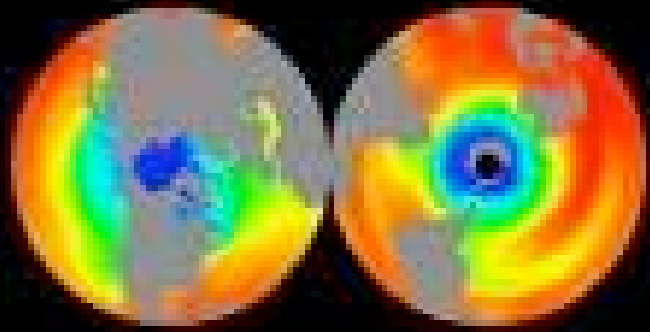
- **Programming:** VHDL, C2Gate?, no cache ✓ Fortran C, CC Memory
- **Compile Time:** Place/Route overnight ✓ Personalities
- **Cost:** HPC addition

✓ Convey focus

# Applications



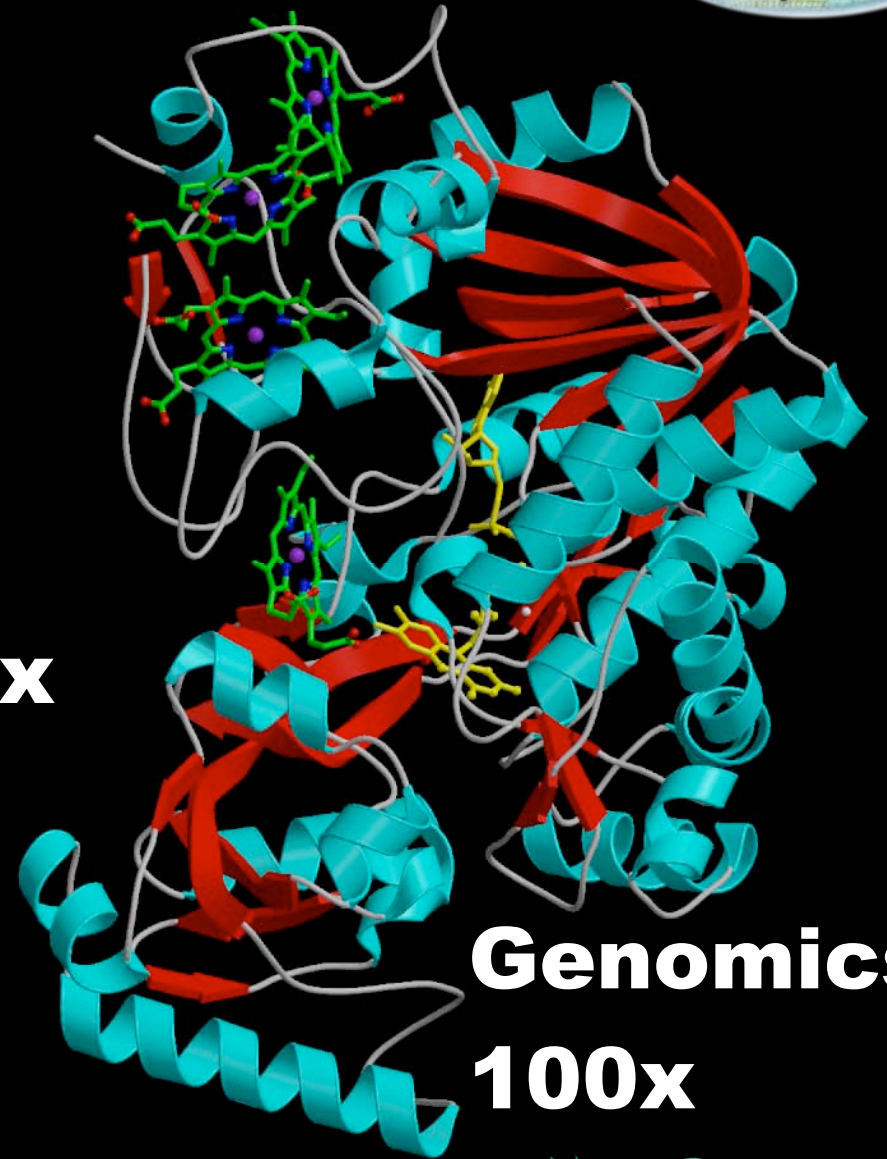
**Weather/Climate-7x**



**Molecular Dynamics-8x**

**Equation Solution-10x**

$$[A]\{x\} = \{b\}$$



**Genomics  
100x**

# FASTA Sequencing Code for Human DNA

- **FASTA**: <http://fasta.bioch.virginia.edu>
- **search34** code & Cray **Smith-Waterman** core
- **Human Genome Data: 4GB compressed**  
**3685 searches (MPI on ORNL Cray XD1)**



Alignment of ACGAACCCTTGC and ACGTATGC

	0	A	C	G	T	A	T	G	C
0	0	0	0	0	0	0	0	0	0
A	0	2	0	0	0	2	0	0	0
C	0	0	4	2	1	0	1	0	2
G	0	0	2	6	4	3	2	3	1
A	0	2	1	4	5	6	4	3	2
A	0	2	1	3	3	7	5	4	3
C	0	2	4	2	2	5	6	4	6
C	0	0	2	3	1	4	4	5	6
C	0	0	2	1	2	3	3	3	7
T	0	0	0	1	3	2	5	3	5
T	0	0	0	0	3	2	4	4	4
G	0	0	0	2	1	2	2	6	4
C	0	0	2	0	1	0	1	4	6

Final alignment

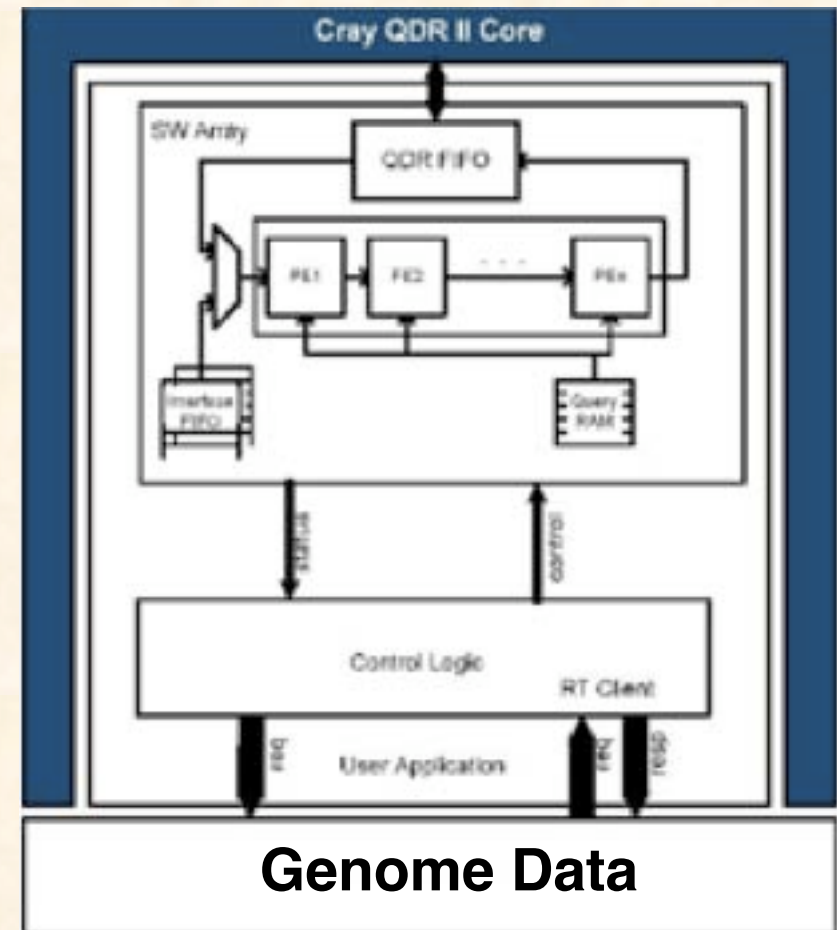
A	C	G	A	A	C	C	C	T	T	G	C
A	C	G	T	A	-	-	-	-	T	G	C

# Smith-Waterman Pipeline Algorithm

## Parallel Score Calculation

		Query Sequence						
		0	A	C	G	T	...	C
Database Sequence	0	0	0	0	0	0	0	0
	C	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	0
	T	0	0	0	0	0	0	PE N
	⋮	0	0	0	0	0	PE	↓
	T	0	0	0	0	PE 4	↓	
	A	0	0	0	PE 3	↓		
	A	0	0	PE 2	↓			
	G	0	PE 1	↓				
	C	0	↓					
A	0							

## Overall Algorithm



# Smith-Waterman Scoring Algorithm

Query Sequence

Database  
Sequence

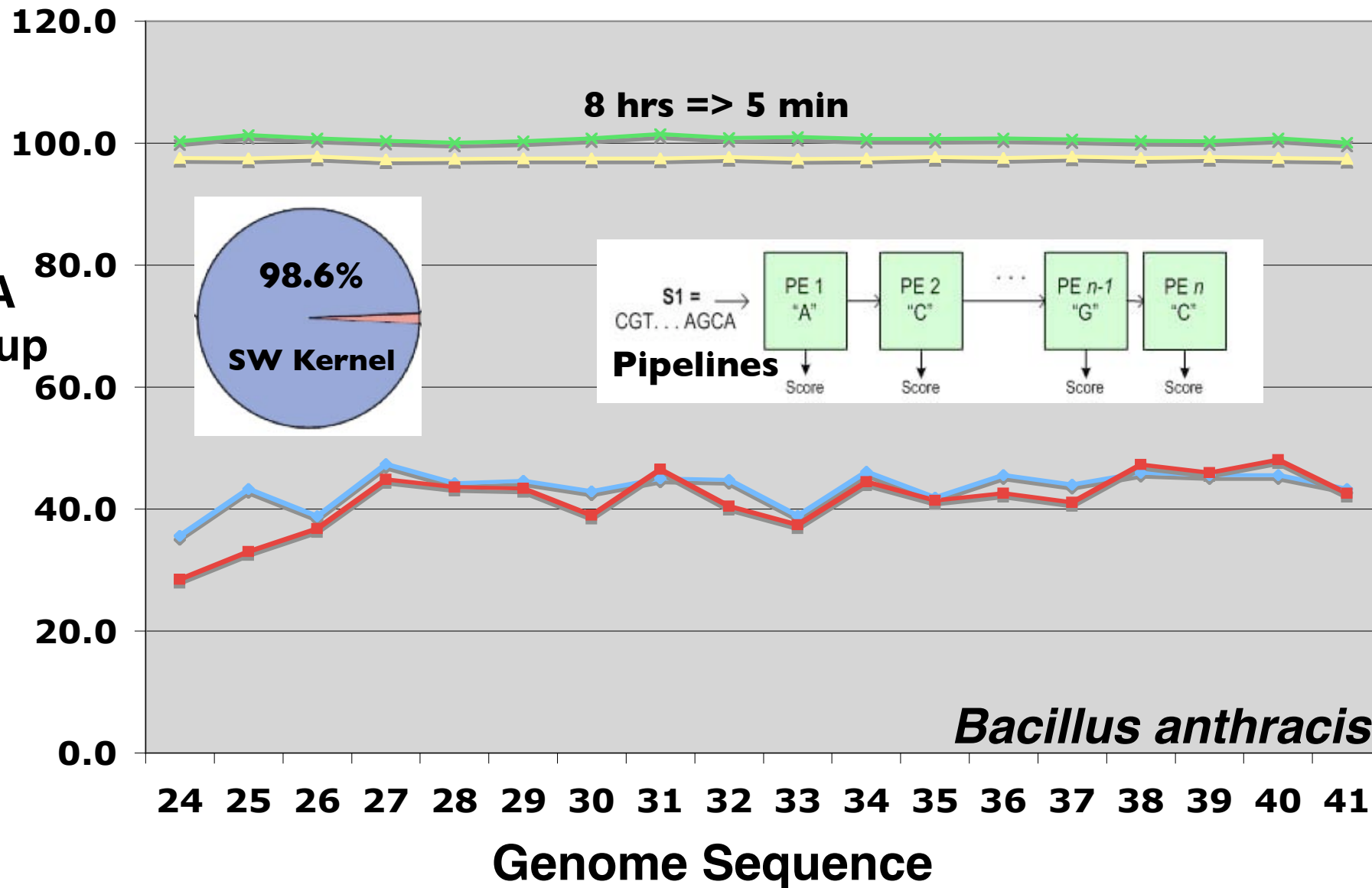
	0	A	C	G	T	...	C
0	0	0	0	0	0	0	0
A	0	2	0	0	0	2	0
C	0	0	4	2	1	0	2
G	0	0	2	6			
A	0						
A	0						
C	0						
...	0						
G	0						

1. Initialize row & column 1 to 0
2. Score matches from upper left
3. Add to above-left score ( $2+4=6$ )



# 100x\* Speedup for Human DNA Sequencing

8k w/align 16k w/align 8k w/o align 16k w/o align



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

# Solution Time on 150 2.2 GHz Opterons @NRL

Job ID User Queue Jobname SessID NDS TSK Memory Time S Time

**Solution Time**

```
-----  
136264 stren compute run_001_op 14310 1 4 -- 900:0 R 745:5 (63-44) 19 seq to go => 1066 hours  
136265 stren compute run_050_op 14320 1 4 -- 900:0 R 745:5 (3150-3128) 22 seq to go => 1144 hours  
136266 stren compute run_100_op 14335 1 4 -- 900:0 R 745:5 (6300-6278) 22 seq to go => 1144 hours  
136267 stren compute run_150_op 14555 1 4 -- 900:0 R 745:5 (9450-9428) 22 seq to go => 1144 hours
```

**Opteron Solution time = 1,144 Hours = 47.66 days => 6 weeks**

```
stren.c494n6% grep ">>" run_001_opteron.out | tail -1 44>>>chrX_016k_seq000044 - 16350 nt  
stren.c494n6% grep ">>" run_050_opteron.out | tail -1 41>>>chrX_016k_seq003128 - 16350 nt  
stren.c494n6% grep ">>" run_100_opteron.out | tail -1 41>>>chrX_016k_seq006278 - 16350 nt  
stren.c494n6% grep ">>" run_150_opteron.out | tail -1 41>>>chrX_016k_seq009428 - 16350 nt
```

Near completion thru 63 total sequences:

```
stren.c494n6% grep ">" chrX_16k_run001.fa | tail -1 >chrX_016k_seq000063  
stren.c494n6% grep ">" chrX_16k_run050.fa | tail -1 >chrX_016k_seq003150  
stren.c494n6% grep ">" chrX_16k_run100.fa | tail -1 >chrX_016k_seq006300  
stren.c494n6% grep ">" chrX_16k_run150.fa | tail -1 >chrX_016k_seq009450
```

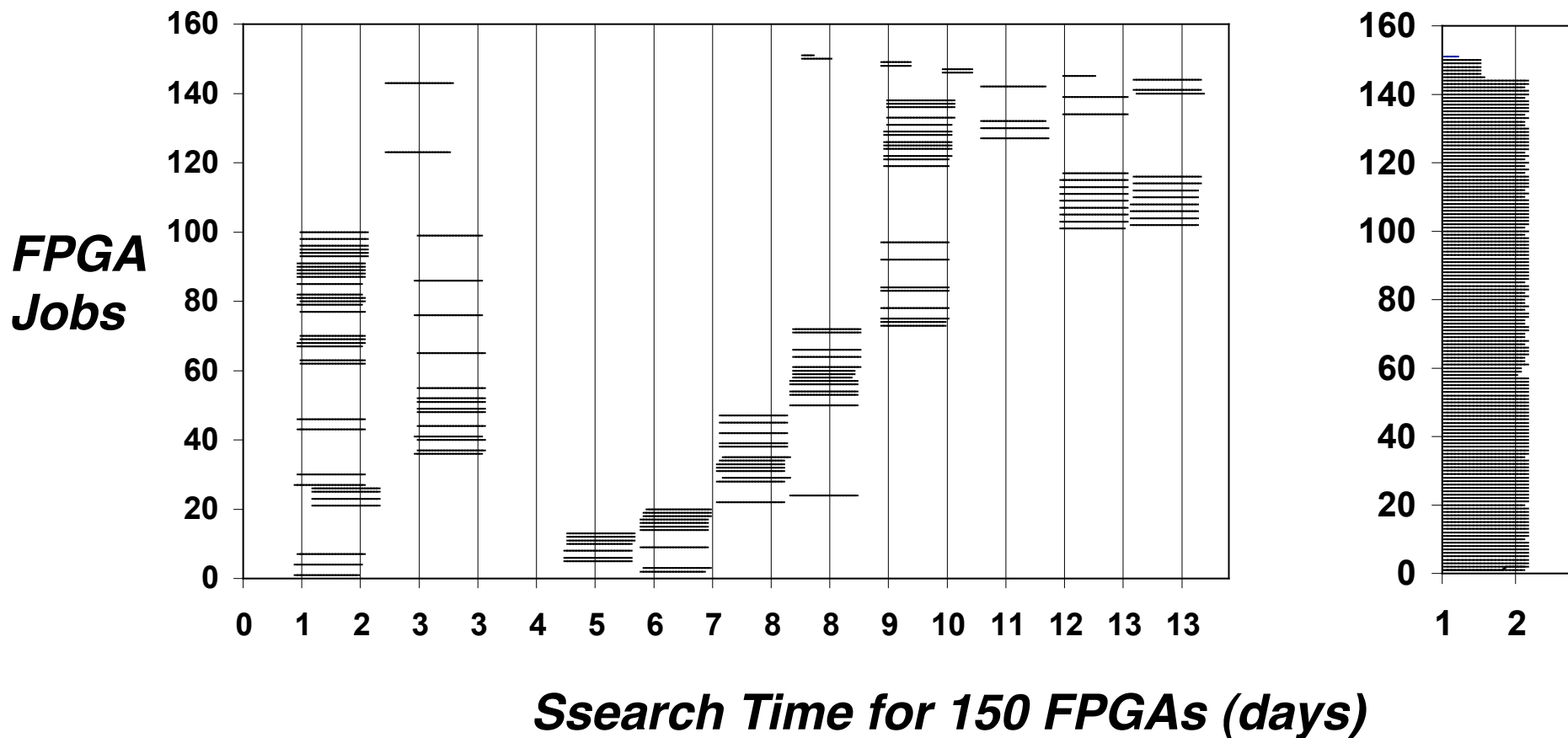
**FPGA Solution time = 24 hrs ~ 48X speedup over Opteron**  
**but dominated by Opteron I/O**

# DNA Sequence\* Time on 150 FPGAs

\* Human-Mouse DNA Compare (FASTA)

*“Non-dedicated” FPGAs*

*Dedicated FPGAs*



# DNA Sequencing: Speed\* on 150 FPGAs

\* State-of-the-art: **G**iga **C**ell **U**dates **P**er **S**econd (**GCUPS**)

❖ **DNA Characters:** Human = 155 million, Mouse = 165 million

**Total Compares** =  $155\text{M} \times 165\text{M} \times 106^2 \times 2 = 51 \times 10^{15}$  Cell Updates

❖ **Sequential FPGA** ==> 138 days (11,923,200 secs) ==> 4.3 TCUPS  
( $51 \times 10^{15} / 11,923,200$ )

❖ **Parallel (actual)** ==> 12.9 days (1,114,560 secs) ==> 46 TCUPS

❖ **Parallel (dedicated)** ==> 1 day (86,400 secs) ==> 605 TCUPS

## I/O Bottleneck: FPGA stops for Opteron Writes

### Remedy: Replace N writes by one binary write

**Change:** do 100 i=1,n  
write(6,110) x(i),y(i),z(i)  
100 continue  
110 format (1pe13.5, 1pe13.5, 1pe13.5)

**To:** write(format\_string,200) '(' ,n,'(1pe13.5,1pe13.5,1pe13.5\))'  
200 format (a1,i3,a20)  
write(6,201) (x(i),y(i),z(i),i=1,n)  
201 format (format\_string)

**Or:** write formatted data to large character buffer in //  
& copy buffer to disk in one binary write.

# Up to 10x Speedup by reduced I/O

(all alignment output options benefit)

❖ **DNA Characters:** Human = 155 million, Mouse = 165 million

**Total Compares** =  $155\text{M} \times 165\text{M} \times 106^2 \times 2 = 51 \times 10^{15}$  Cell Updates

❖ **Sequential FPGA:** 138 days => 13.8 days\* => 43 TCUPS

❖ **Parallel (actual):** 12.9 days => 1.29 days => 460 TCUPS

❖ **Parallel (dedicated):** 1 day => 2.4 hours => 6 PCUPS

\* with 10X Speedup

# Speedup on 150 FPGAs\*

1 Opteron ==> 20 years (240 mos.)

1 FPGA ==> 5 months

150 Opterons ==> 6 weeks ✓

150 FPGAs ==> 1 day ==> 49X speedup - Virtex2 ✓

10X I/O Speedup {  
    ==> 12 hours ==> 98X speedup - Virtex4  
    ==> 2.4 hours ==> 490X speedup - Virtex2  
    ==> 1.2 hours ==> 980X speedup - Virtex4

\* Compared to Cray XD1's 2.2 GHz Opteron

# Summary



- ***FPGAs increasingly attractive to HPC***
  - **Low power**, faster speed, telecom spinoff (stable)
  - Downsides being addressed (coding, memory speed)
  - New vendor options: *Cray, Convey,...*
  - **100X Genomics Speedup best**: scalable to 150 FPGAs
  - Streamed I/O offers additional 10X speedup
- ***Accelerators key to bring HPC to the “next level”***

Acknowledgment: This is a work of the U.S Government (public domain) supported by the Office of Science, U.S. Department of Energy Contract DE-AC05-00OR22725

The authors thank the US Naval Research Laboratory for access to the 150 FPGA Cray XD1



# THANK YOU!

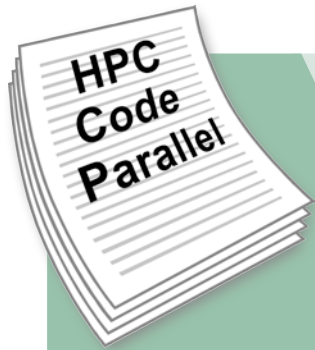


## Contact

**Olaf O. Storaasli**  
Google Olaf ORNL



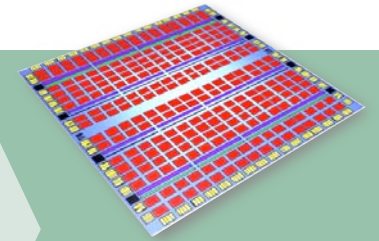
# Weather-Climate code port to FPGAs



Profile-Develop HLL

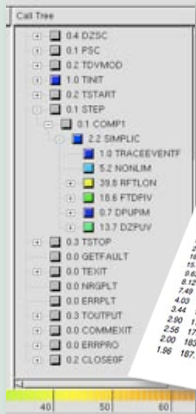


HLL compiler  
CHiMPS, Mitrion  
(FPGA Tools Inside)



FPGA speedup

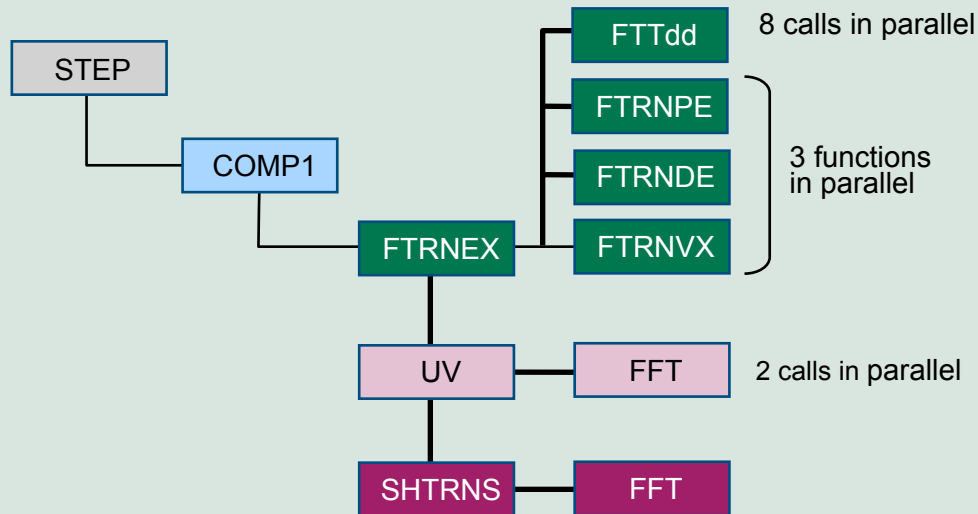
## Profile



Call profile:  
Data sample counts as 0.01 seconds.

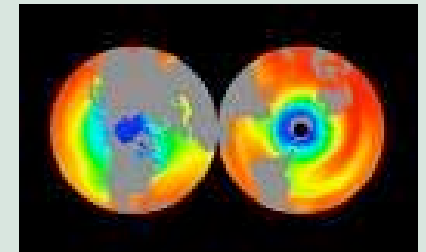
% cumulative	self	self	total	%		
time	seconds	calls	calls	local name		
20.82	40.07	11	3.64	5.23	errand...	
18.82	26.24	129	0.28	0.28	vpassm...	
13.40	106.00	29.66	10	2.97	2.97	fitm...
8.92	124.65	18.55	11	1.69	2.95	shrms...
8.12	140.19	15.64	11	1.42	1.74	shnrc...
7.49	154.48	14.43	33	0.44	0.44	ms9d...
4.93	162.39	7.77	1	7.77	192.64	MAIN...
3.44	169.01	6.62	1	6.62	6.74	calp...
2.90	174.59	5.58	1	5.58	8.12	pd...
2.28	179.52	4.93	10	0.49	10.22	advect...
2.00	183.37	3.85	10	0.39	0.39	ms9a...
1.86	187.15	3.78	356	0.01	0.01	ordlog...

## Find parallelism: 80% FFTs



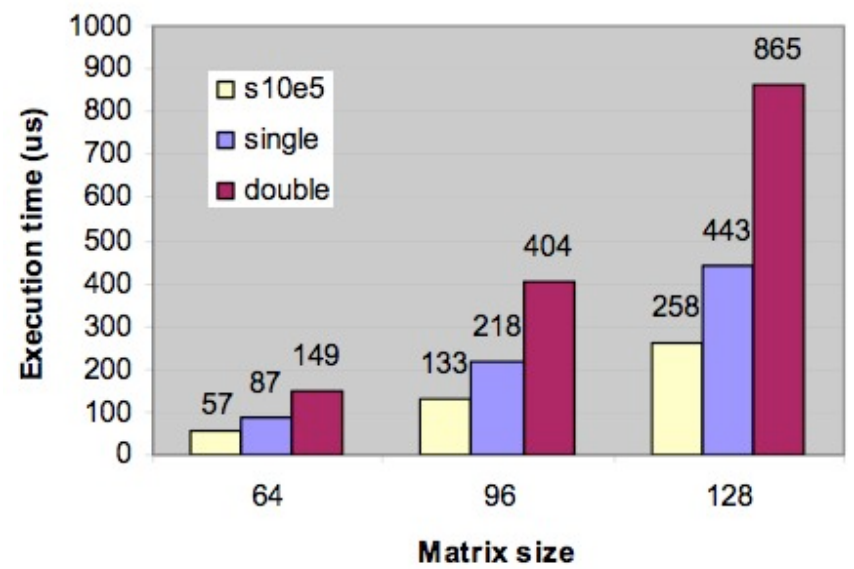
## Goal

More GF/\$ GF/Watt



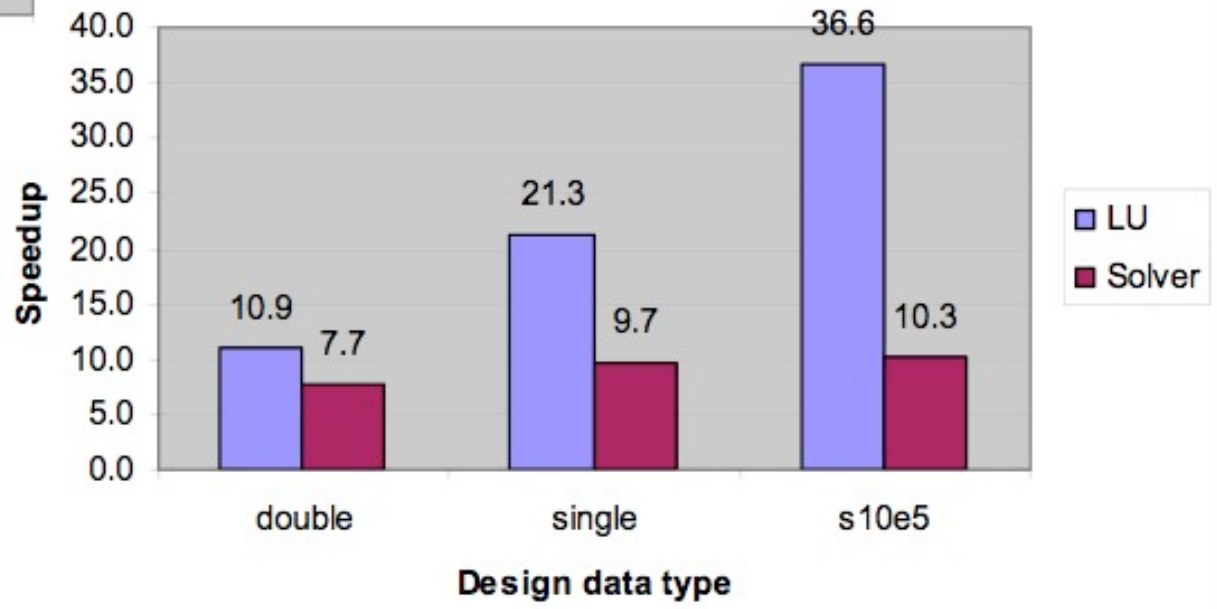
7X speedup

# 37x\* LU Decomposition FPGA Speedup 10x for Matrix Equation Solver



**Table 6: LU implementation on XC2VP50-7**

Design	Double FP	Single FP	S10e5
PE amount	8	16	32
Max size	128	256	256
Achievable Frequency	120MHz	150MHz	150MHz
Slices	27,005 (57%)	14792 (59%)	14730 (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*  
*(LU dominates calculations)*

**First mixed-precision LU & solver for FPGAs**

\*Virtex-II vs 2.2 GHz Opteron