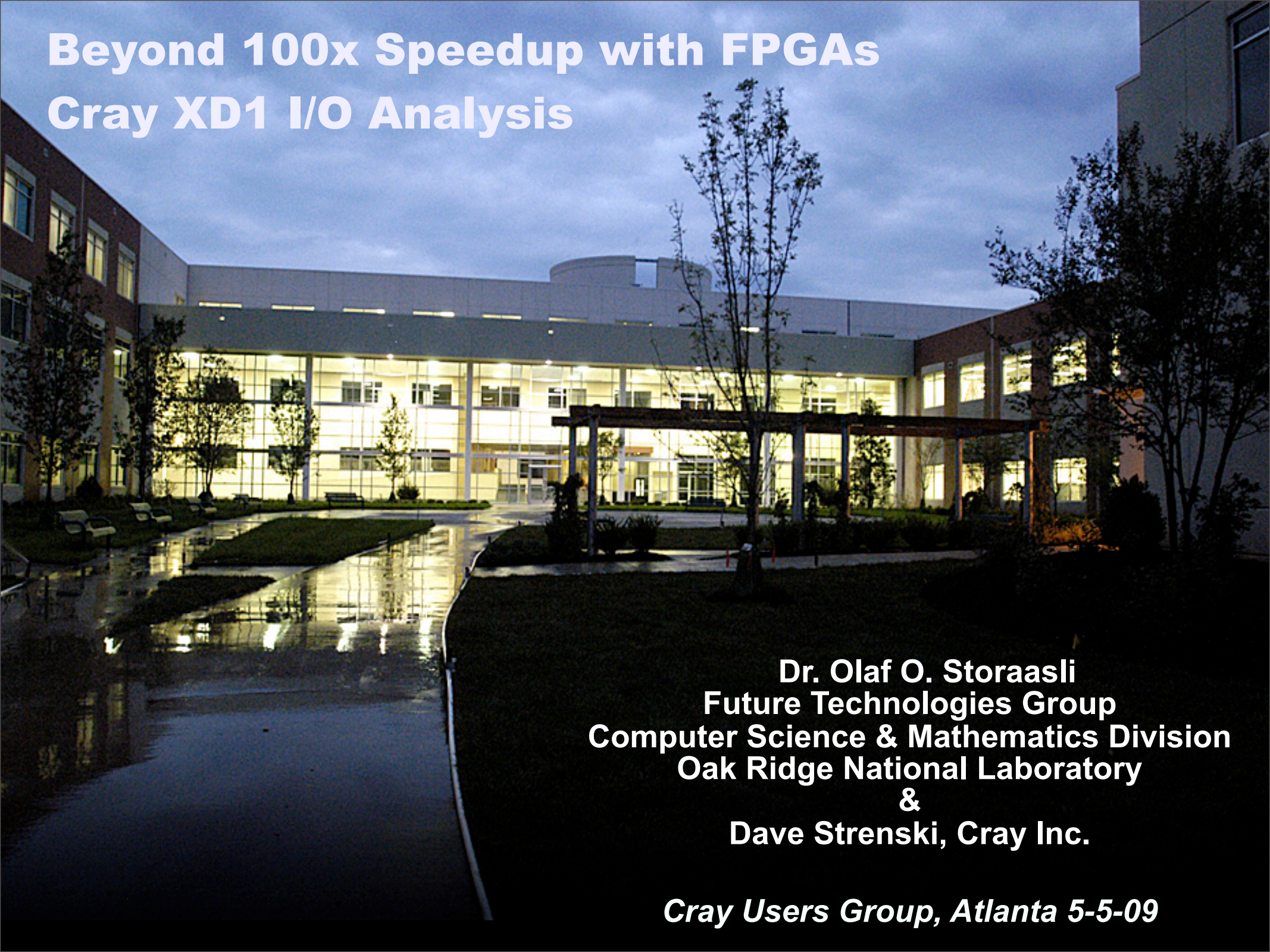
Beyond 100x Speedup with FPGAs
Cray XD1 I/O Analysis



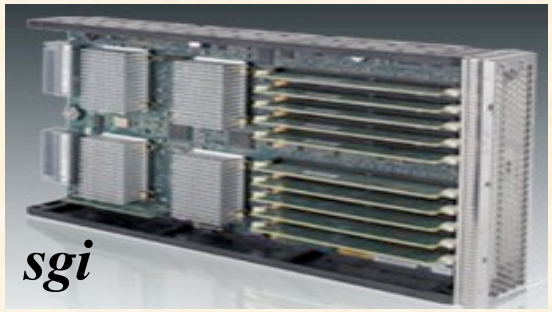
Dr. Olaf O. Storaasli
Future Technologies Group
Computer Science & Mathematics Division
Oak Ridge National Laboratory
&
Dave Strenski, Cray Inc.

Cray Users Group, Atlanta 5-5-09

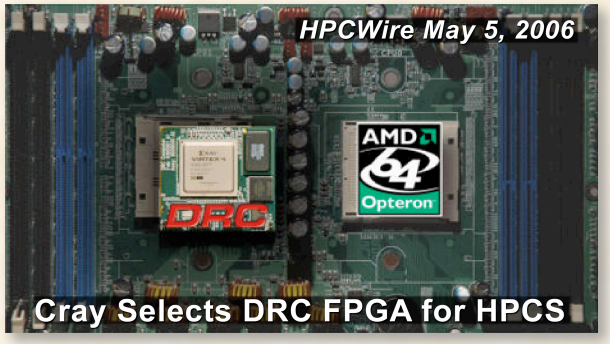
3 FPGA Generations: Moving toward HPC

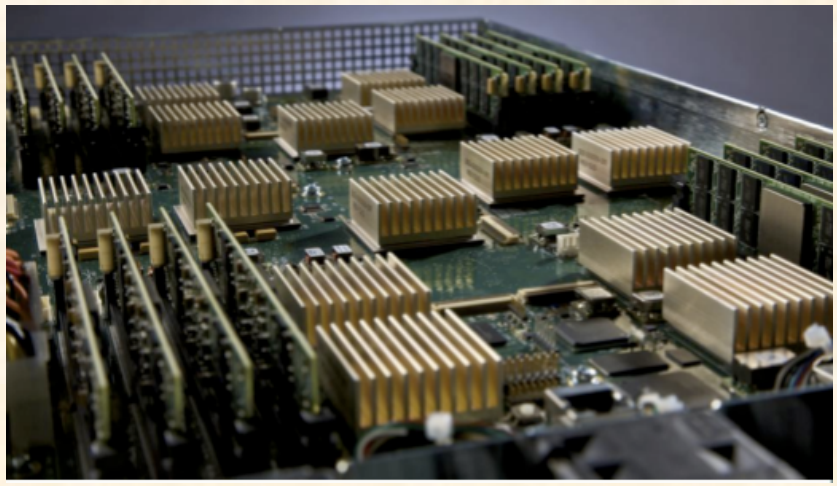
PCI: ANS, DSP => HPEC

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



Socket: Cray XT5h (DRC, XtremeData), Convey





Future Technologies Group ornl

OAK RIDGE National Laboratory

ORNL Cray XD1 with Xlininx 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?

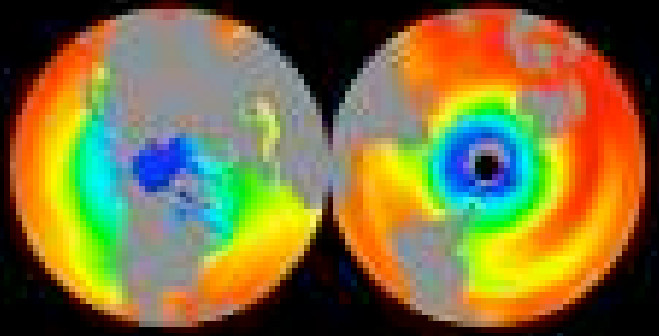
- **Programming**: VHDL, C2Gate?, no cache ✓ Fortran C, CC Memory
- **Compile Time**: Place/Route overnight ✓ Personalities
- **Cost**: HPC addition

✓ Convey focus

Applications

Future Technologies Group

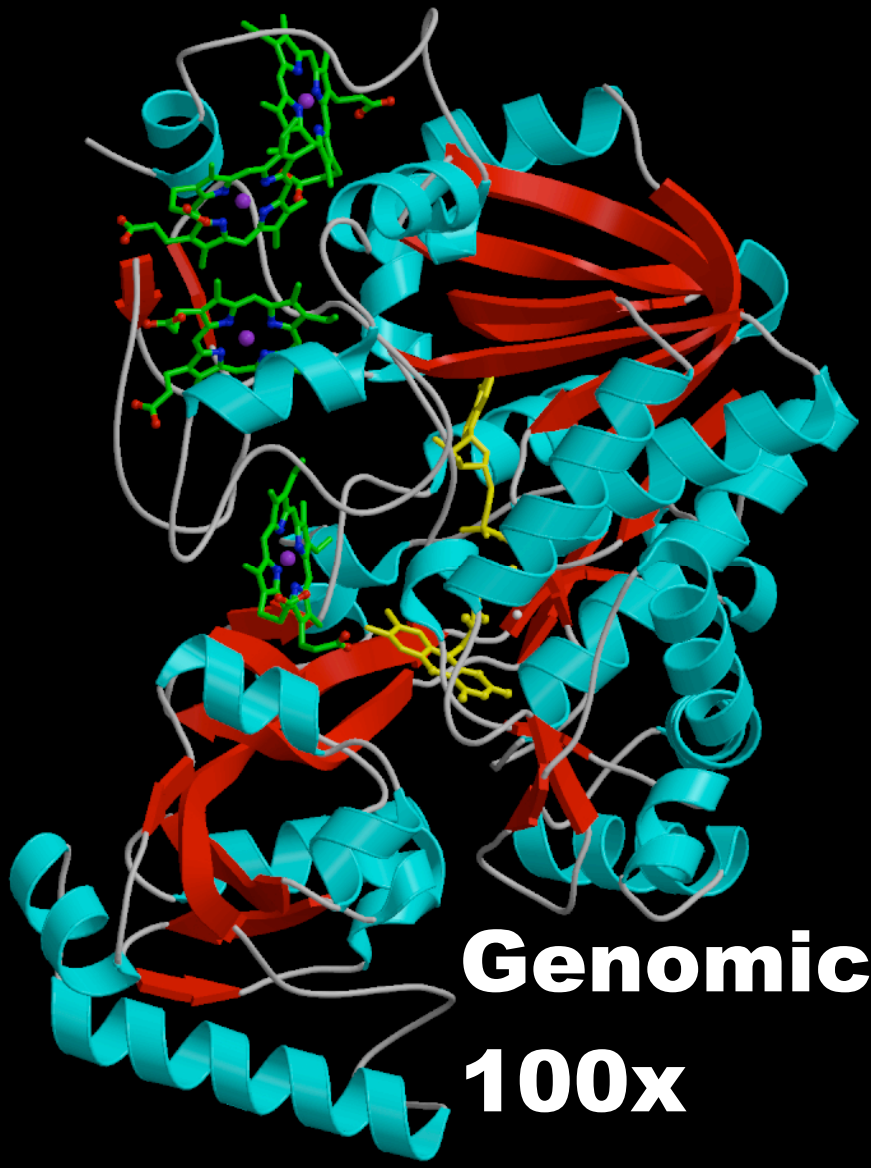
Weather/Climate-7x



Molecular Dynamics-8

Equation Solution-10x

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



Genomics

100x

OAK RIDGE National Laboratory

OAK RIDGE National Laboratory

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

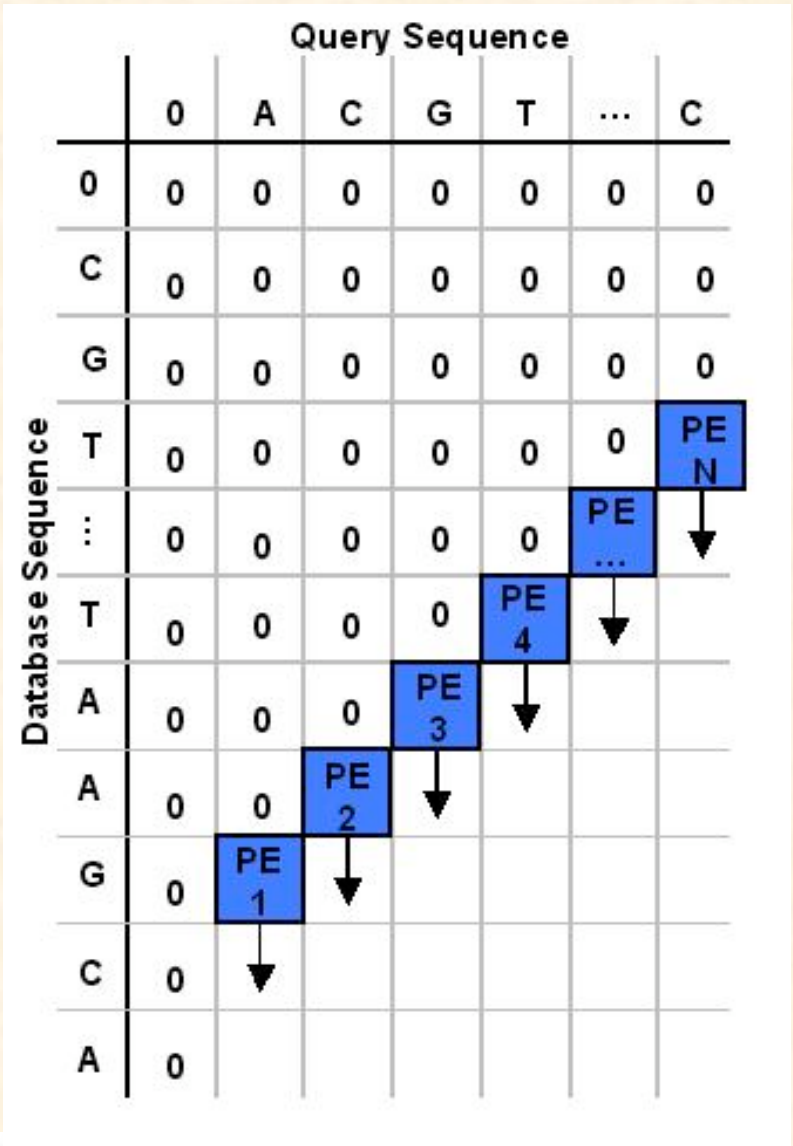
	0	A	C	G	T	A	T	G	C
0	0	0	0	0	0	0	0	0	0
A	0	2	0	0	0	2	0	0	0
C	0	0	4	2	1	0	1	0	2
G	0	0	2	6	4	3	2	3	1
A	0	2	1	4	5	6	4	3	2
A	0	2	1	3	3	7	5	4	3
C	0	2	4	2	2	5	6	4	6
C	0	0	2	3	1	4	4	5	6
C	0	0	2	1	2	3	3	3	7
T	0	0	0	1	3	2	5	3	5
T	0	0	0	0	3	2	4	4	4
G	0	0	0	2	1	2	2	6	4
C	0	0	2	0	1	0	1	4	8

Final alignment

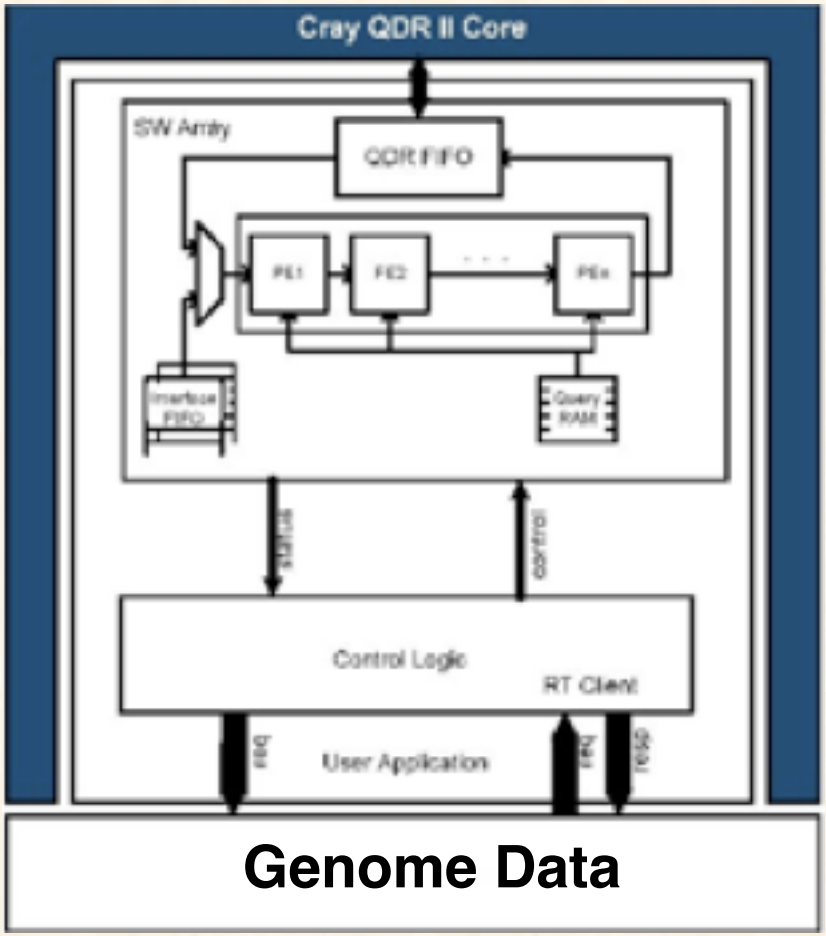
A	C	G	A	A	C	C	C	T	T	G	C
	|	|		|					|	|	|
A	C	G	T	A	-	-	-	-	T	G	C

Smith-Waterman Pipeline Algorithm

Parallel Score Calculation



Overall Algorithm



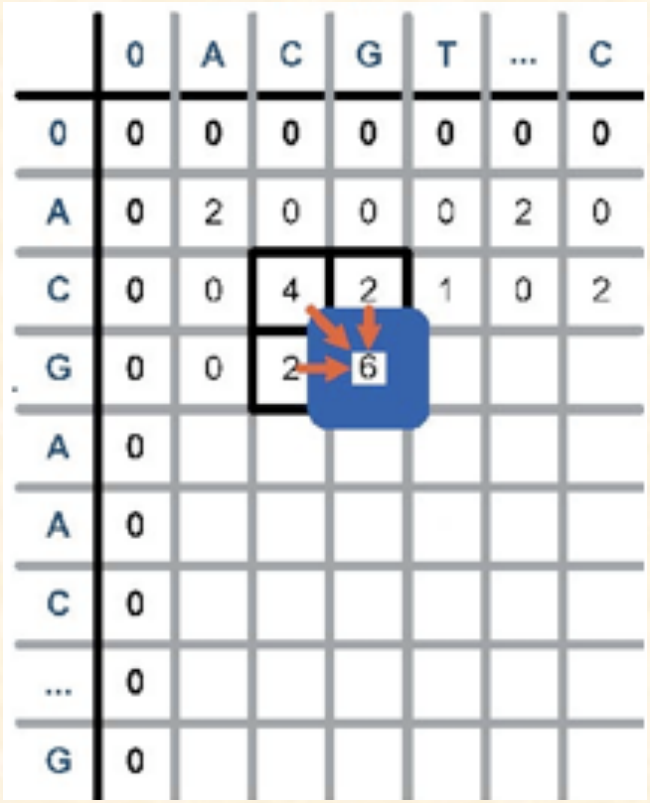
Future Technologies Group ornl

OAK RIDGE National Laboratory

Smith-Waterman Scoring Algorithm

Query Sequence

Database Sequence



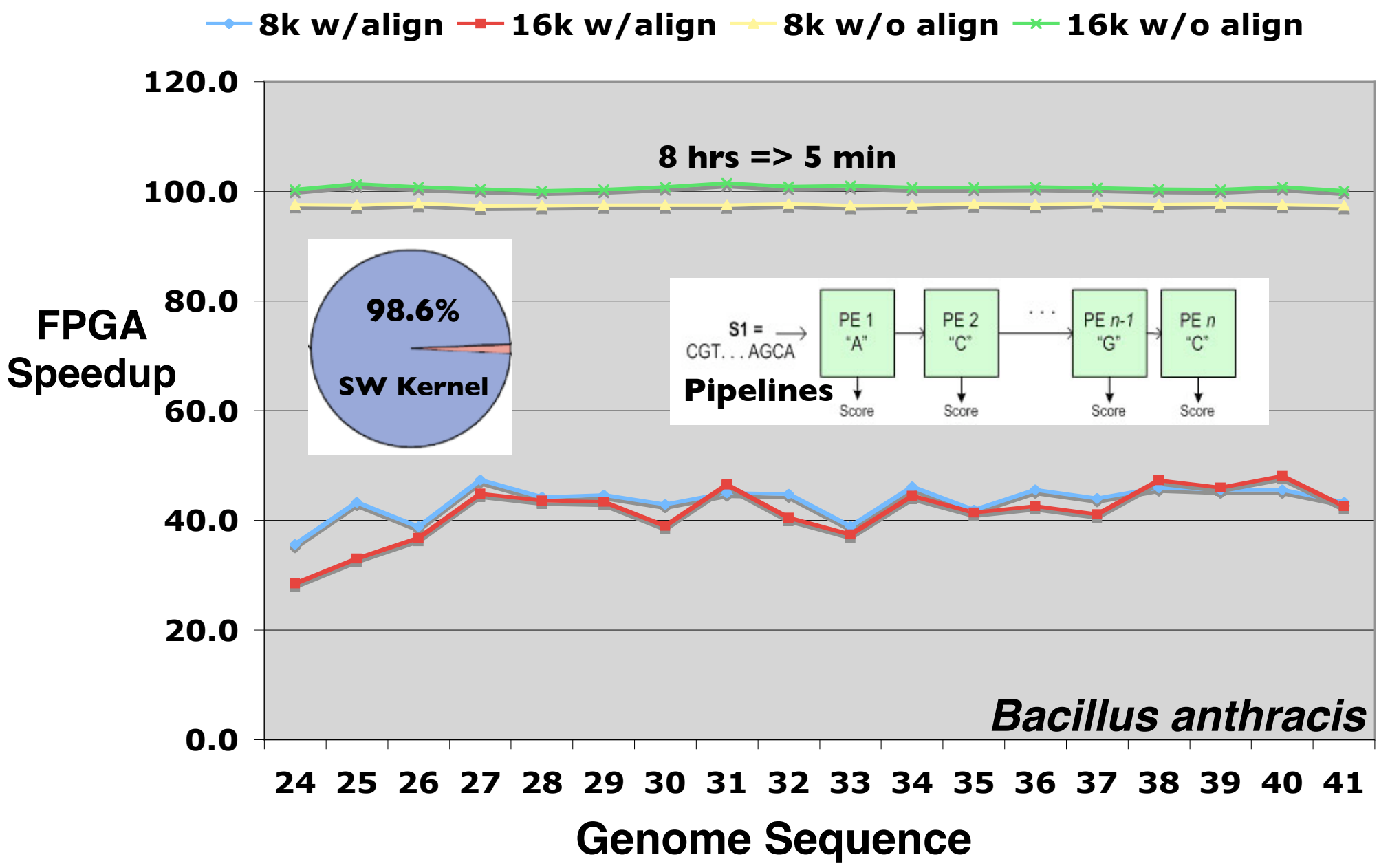
	0	A	C	G	T	...	C
0	0	0	0	0	0	0	0
A	0	2	0	0	0	2	0
C	0	0	4	2	1	0	2
G	0	0	2	6			
A	0						
A	0						
C	0						
...	0						
G	0						

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

Future Technologies Group ornl

OAK RIDGE National Laboratory

100x* Speedup for Human DNA Sequencing



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

OAK RIDGE National Laboratory

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	1066 hours
136265	stren	compute	run_050_op	14320	1	4	--	900:0	R	745:5	1144 hours
136266	stren	compute	run_100_op	14335	1	4	--	900:0	R	745:5	1144 hours
136267	stren	compute	run_150_op	14555	1	4	--	900:0	R	745:5	1144 hours

(63-44) 19 seq to go =>
(3150-3128) 22 seq to go =>
(6300-6278) 22 seq to go =>
(9450-9428) 22 seq to go =>

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
```

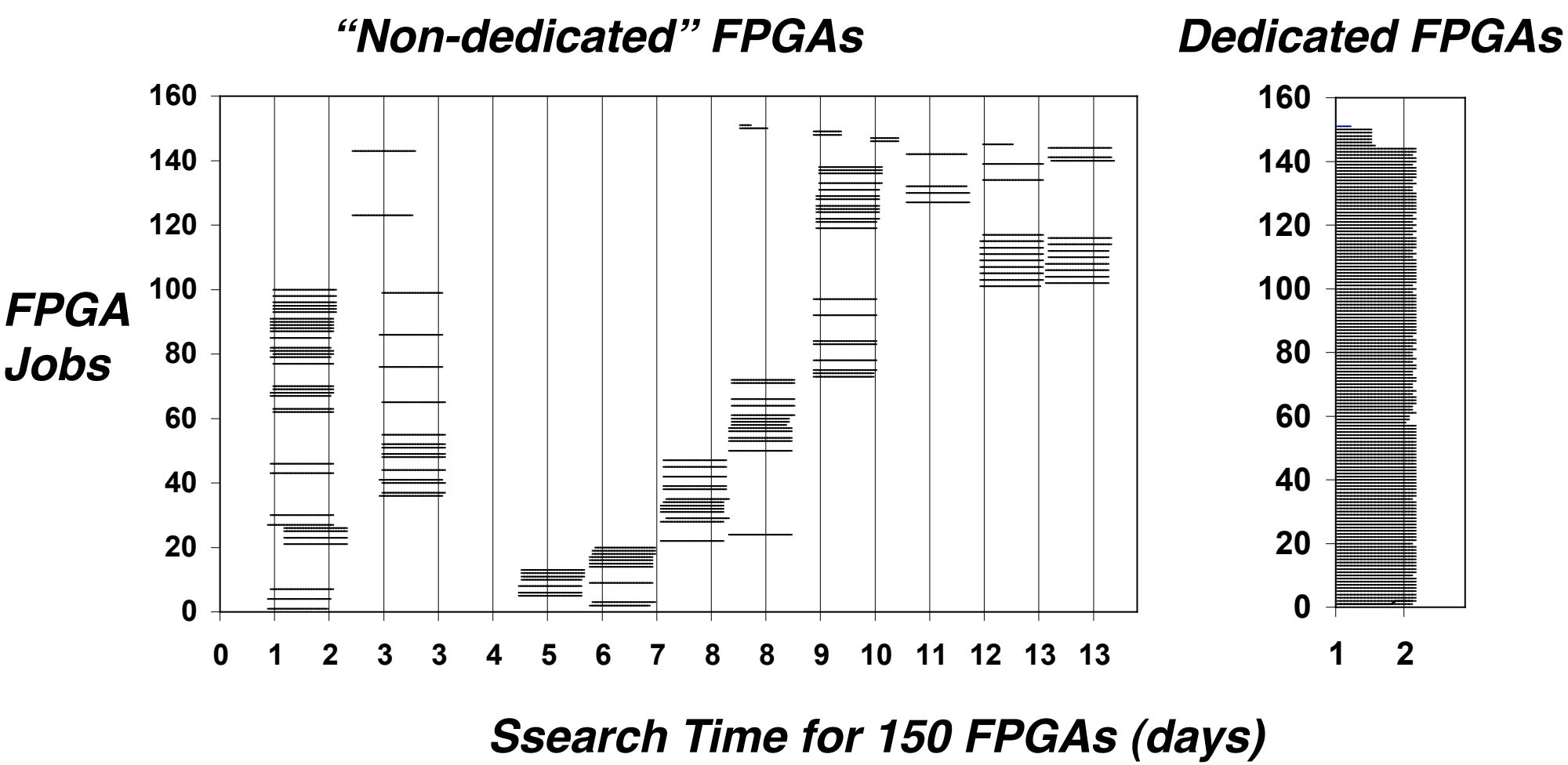
Near completion thru 63 total sequences:

```
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)



DNA Sequencing: Speed* on 150 FPGAs

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

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

Total Compares = 155M x 165M x 106^2 x 2 = $51x10^{15}$ Cell Updates

- Sequential FPGA ==> 138 days (11,923,200 secs) **==> 4.3 TCUPS**
  (**$51x10^{15}$/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 = 155M x 165M x 106^2 x 2 = 51×10^{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 ✓

==> 12 hours ==> 98X speedup - *Virtex4*

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

* Compared to Cray XD1's 2.2 GHz Opteron

Summary

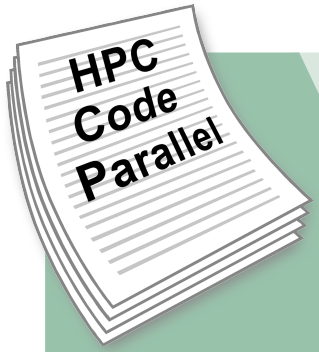
- *FPGAs increasingly attractive to HPC*
  - **Low power, faster speed, telecom spinoff (stable)**
  - **Downsides being addressed (coding, memory speed)**
  - **New vendor options:** *Cray, Convey,…*
  - **100X Genomics Speedup best: scalable to 150 FPGAs**
  - **Streamed I/O offers additional 10X speedup**
- ***Accelerators may 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!
Future Technologies Group
ornl
Contact
Olaf O. Storaasli
Google Olaf ORNL
OAK RIDGE
National Laboratory

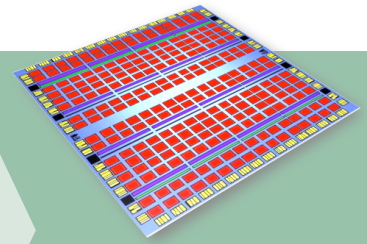
Weather-Climate code port to FPGAs

Future Technologies Group ornl



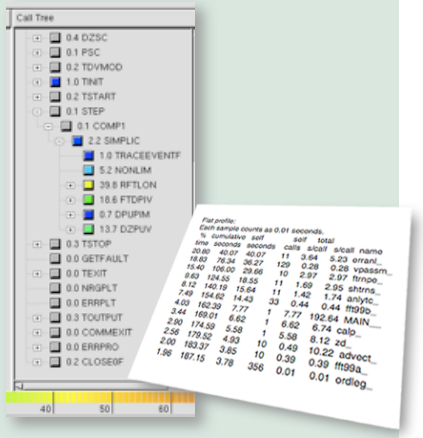
Profile–Develop HLL

HLL compiler
CHiMPS, Mitrion
(FPGA Tools Inside)

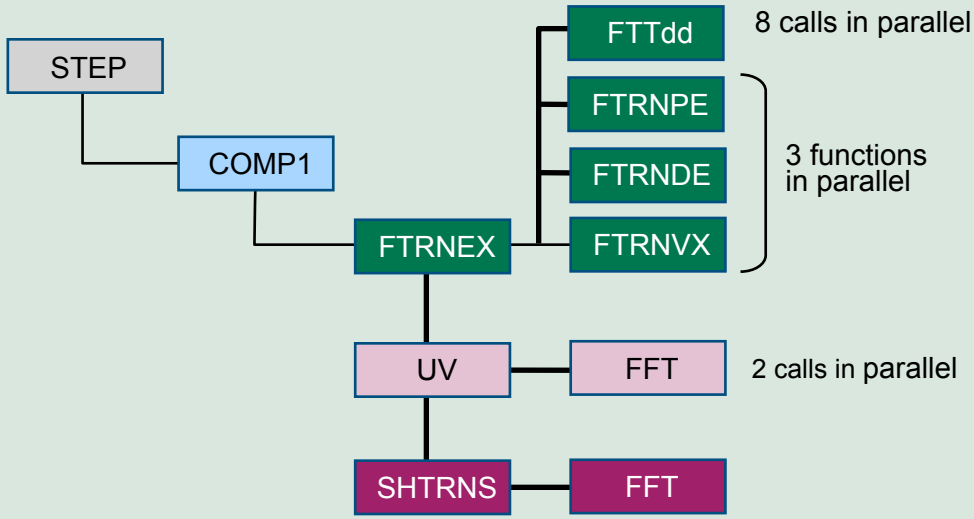


FPGA speedup

Profile

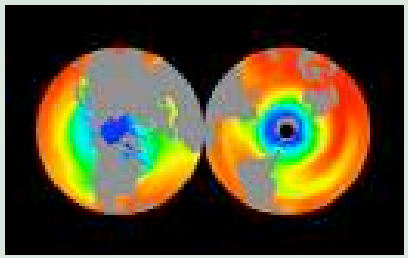


Find parallelism: 80% FFTs



Goal

More GF/$ GF/Watt



7X speedup

OAK RIDGE National Laboratory

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

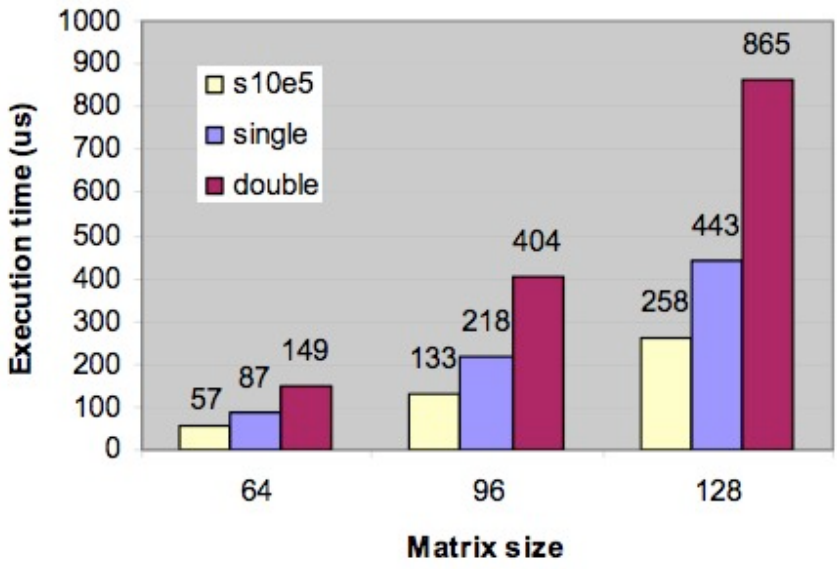
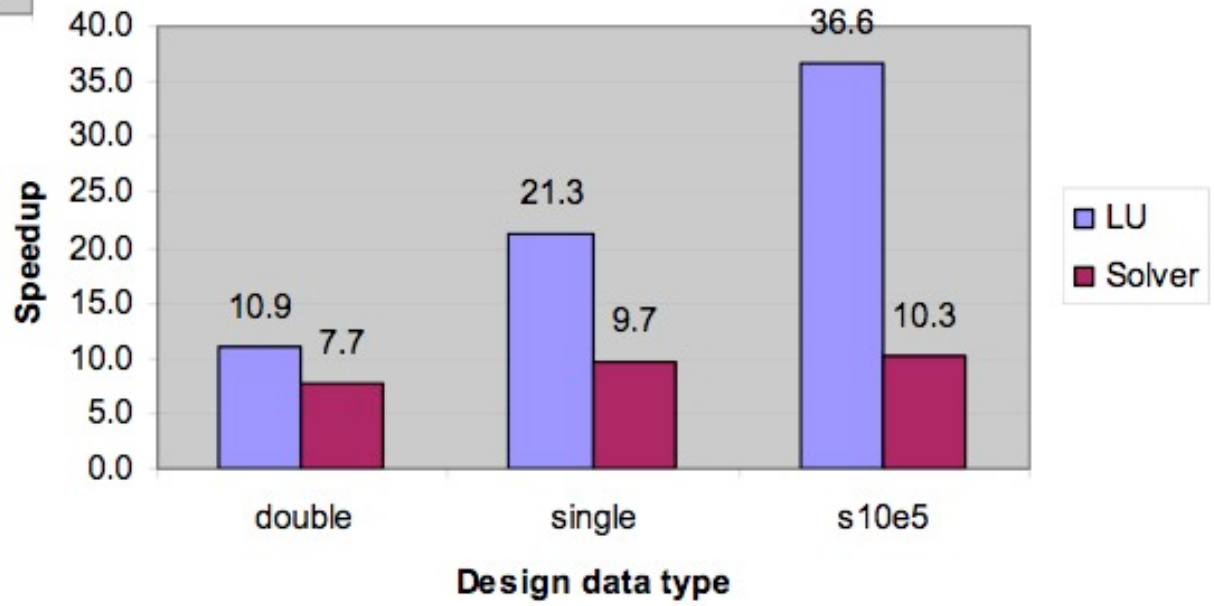


Table 6: LU implementation on XC2VP50-7

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