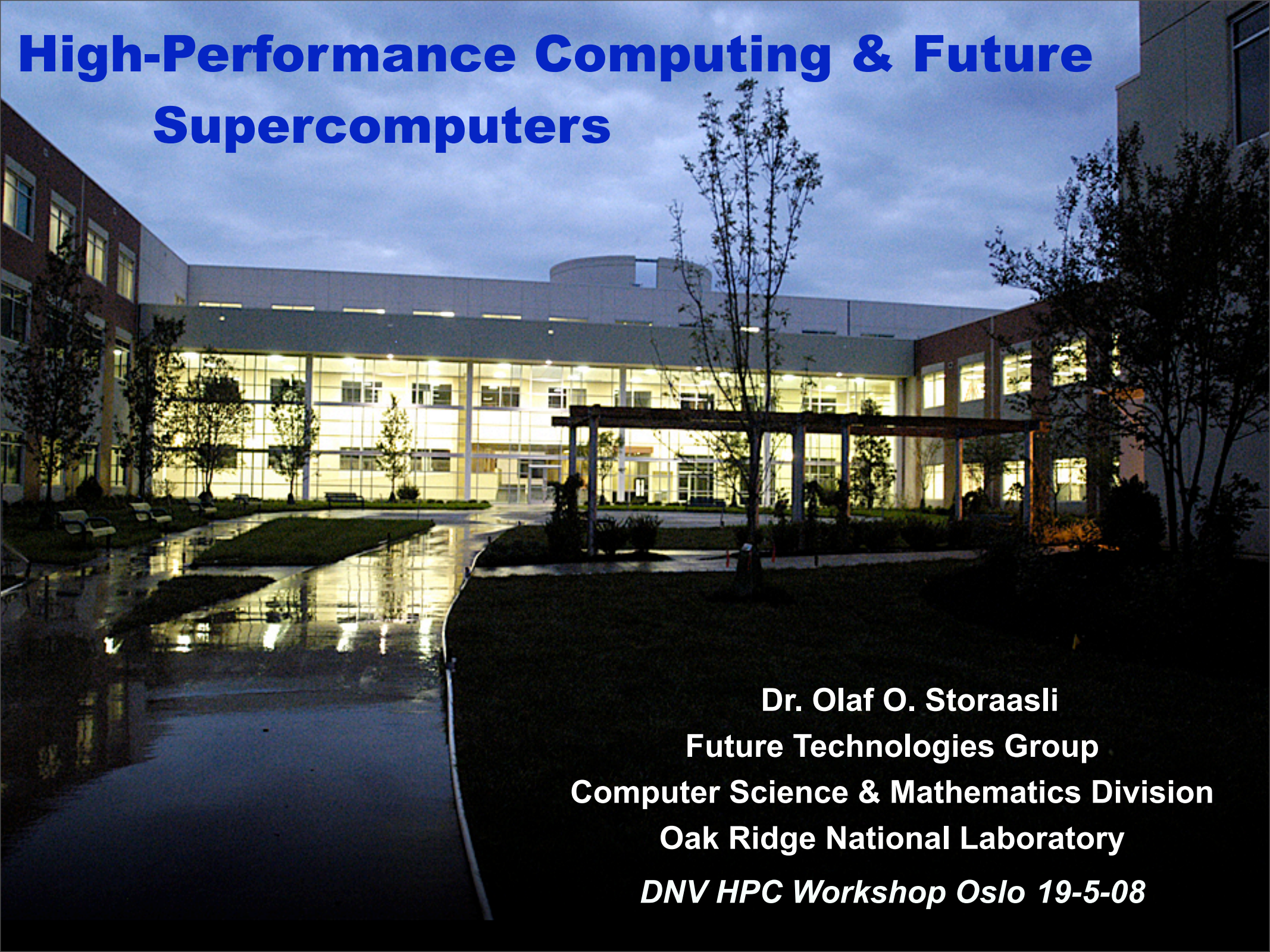
High-Performance Computing & Future Supercomputers



Dr. Olaf O. Storaasli

Future Technologies Group

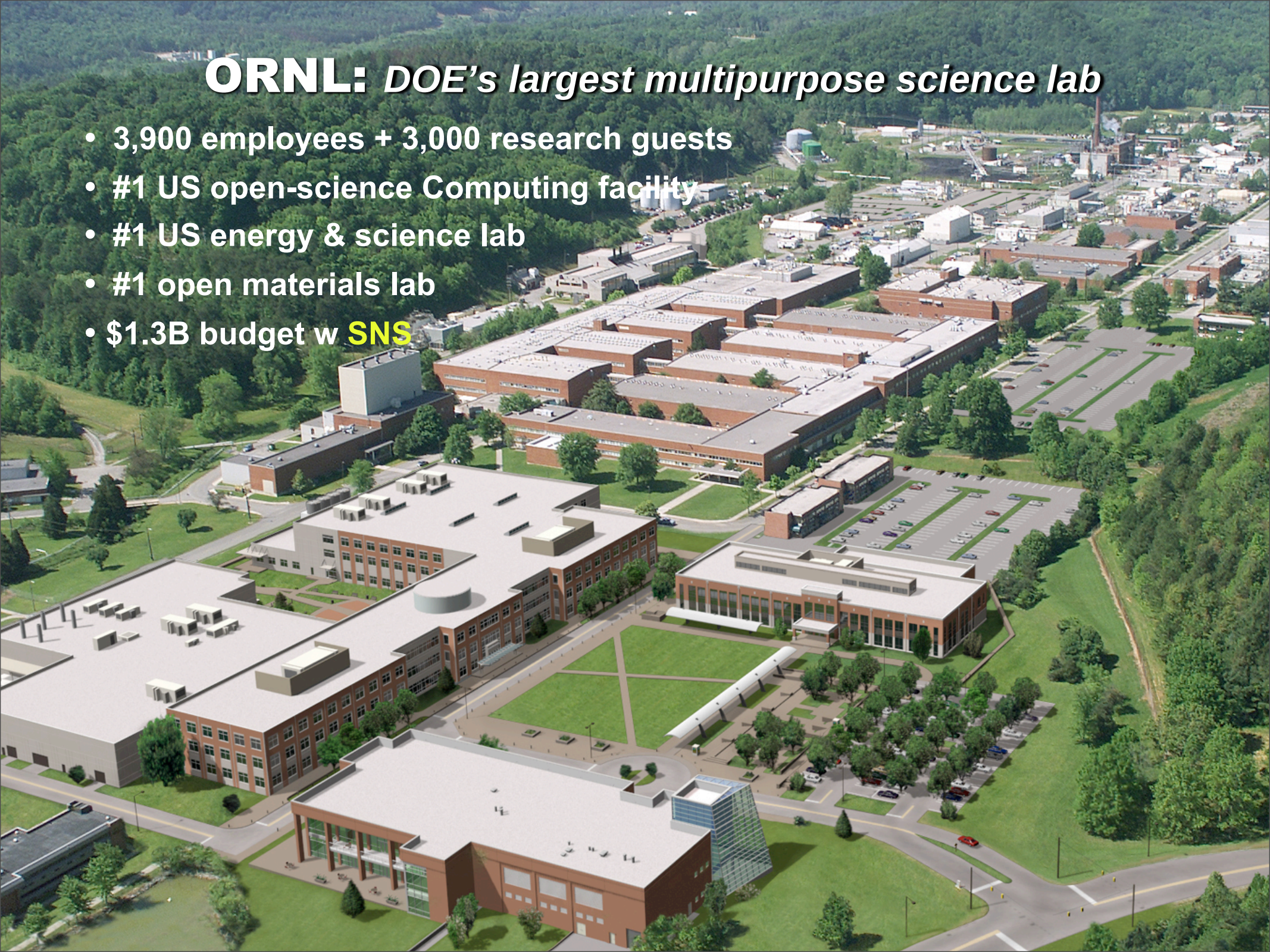
Computer Science & Mathematics Division

Oak Ridge National Laboratory

DNV HPC Workshop Oslo 19-5-08

ORNL “X-10” History

1st Graphite Plutonium Reactor => PNL


ORNL: DOE's largest multipurpose science lab
• 3,900 employees + 3,000 research guests
• #1 US open-science Computing facility
• #1 US energy & science lab
• #1 open materials lab
• $1.3B budget w SNS

ORNL HPC: 102TF => 250TF => 2x1PF

Computer Room 180° Panorama

OAK RIDGE CENTER FOR ADVANCED STUDIES

Future Technologies Group ornl

OAK RIDGE National Laboratory

Context

ORNL Jaguar Supercomputer Advances to Second in the World

System is the world's most powerful for open science

@102TF

Future Technologies Group ornl

OAK RIDGE National Laboratory


Former NASA Langley Engineer
ASTP
Background: FPGAs: NASA => ORNL
Focus: Algorithms => Applications
Goal: Speed Supercomputers with FPGAs

Future Supercomputer Technologies

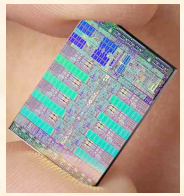
Commodity: 2^n *multi* => *many core*

Special: *El Dorado, Cyclops, PiM*

Accelerators

- **FPGA**: DSP => HPEC => HPC <==

- **Cell**: Sony, Toshiba, IBM



- **GPUs**: => *μ*P

- **Array**: ClearSpeed "niche"

Future Technologies Group ornl

OAK RIDGE National Laboratory

Explore FPGAs for future ORNL HPC

HPC vendors adopting FPGAs

sgi **Virtex4 FPGA blades "accelerate mission-critical applications > 100x"**

CRAY THE SUPERCOMPUTER COMPANY **Steve Scott, CTO HPCWire 24/3/0606**

"*After exhaustive analysis, Cray concluded that, although multi-core commodity processors will deliver some improvement, exploiting parallelism through a variety of processor technologies using scalar, vector, multithreading and* **hardware accelerators** *(e.g.,* **FPGAs** *or ClearSpeed co-processors) creates the* ***greatest opportunity*** *for application acceleration."* => ***Cray XT5h***

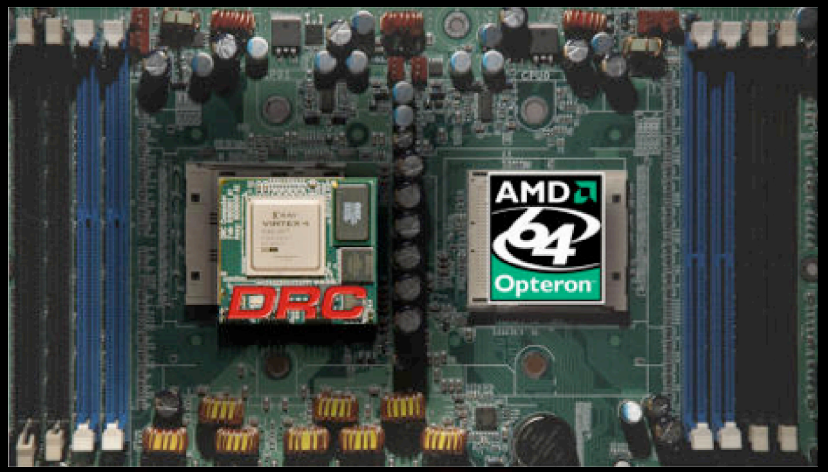
+ HP, SRC, Nallatech, DRC, XtremeData, Inc. COMPUTING REDEFINED **.....**

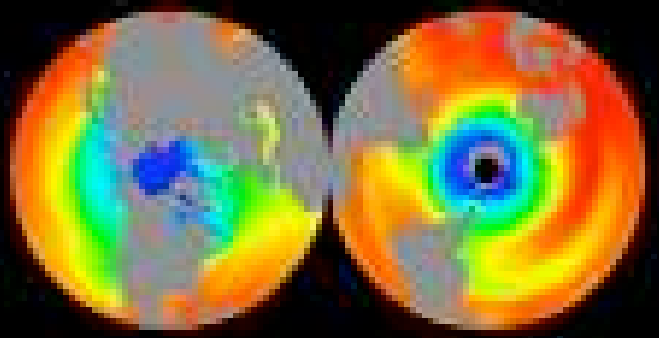
Contents

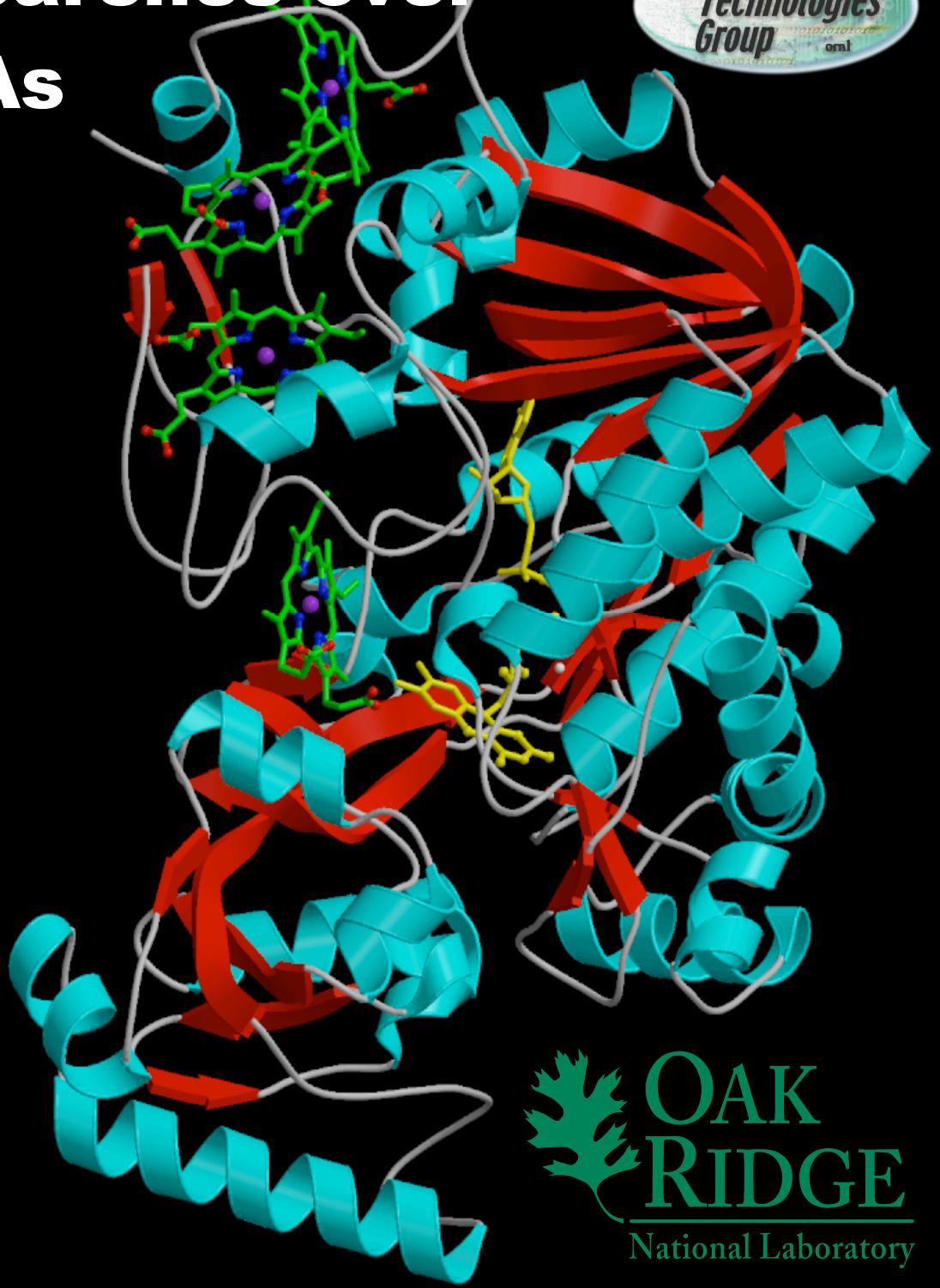
- **Background: Why FPGAs?**
- **ORNL success: FPGA systems, tools and up to 100x speedup**
- **Partners:** XILINX® **Research Lab,** UT**, SRC**
  mitrion CRAY sgi

OAK RIDGE National Laboratory

Speeding Genome Searches over 7350x with 150 FPGAs







Olaf O. Storaasli

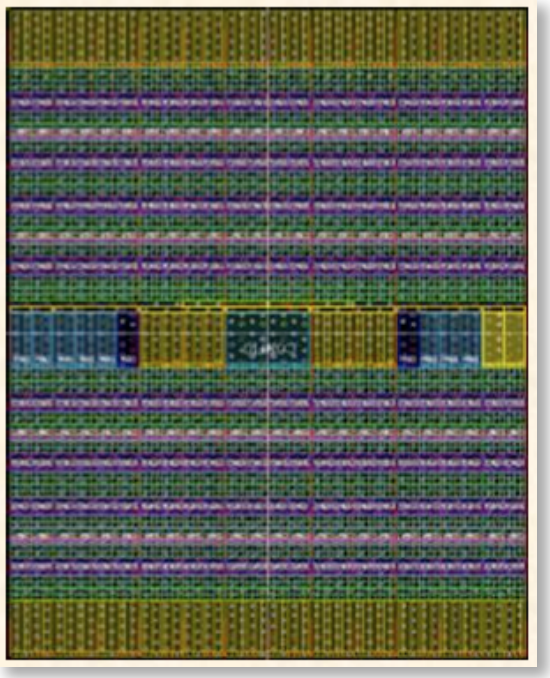
Google **Olaf ORNL**

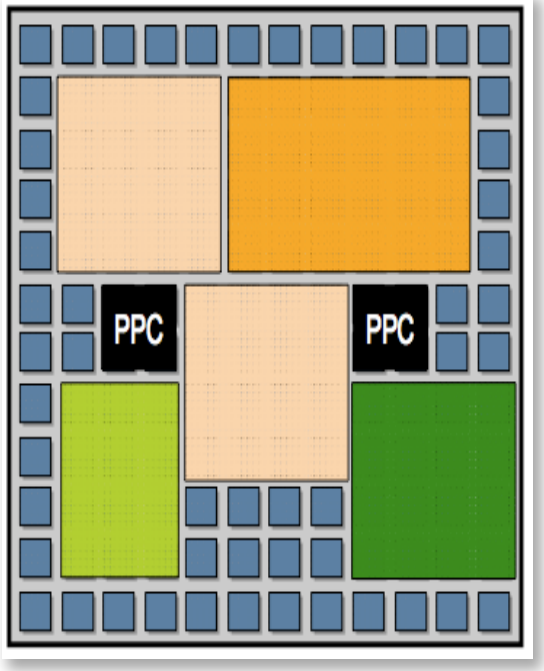
Cray Users Group '08 Helsinki

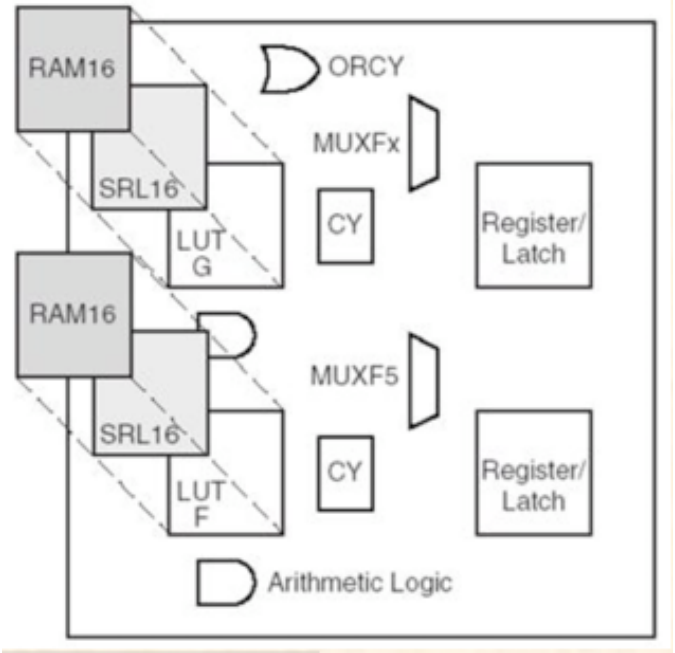
Future Technologies Group ornl

OAK RIDGE National Laboratory

What's an FPGA? Your "custom chip"







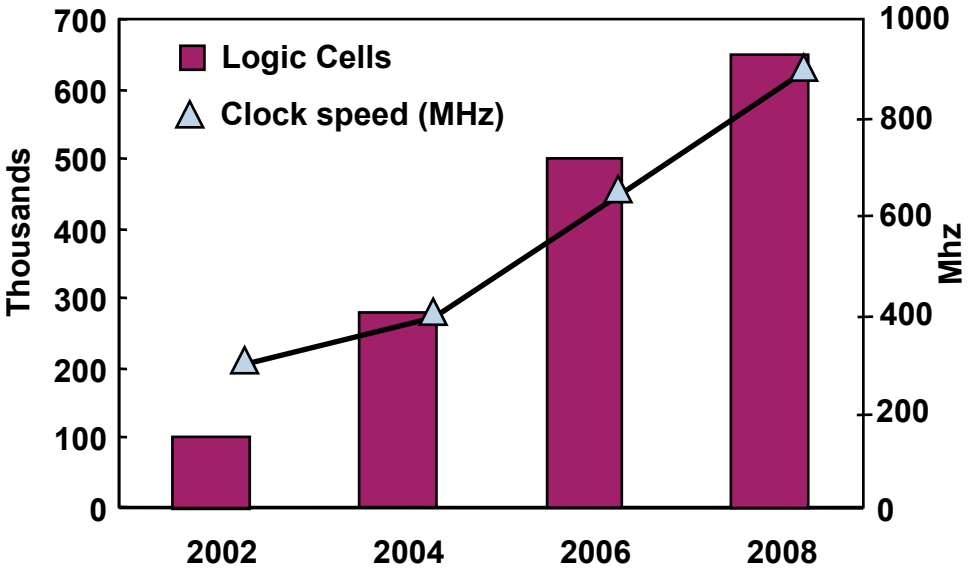
Xilinx Virtex4 FPGA: 89K slices (miniCPUs) FPGA Logic slice

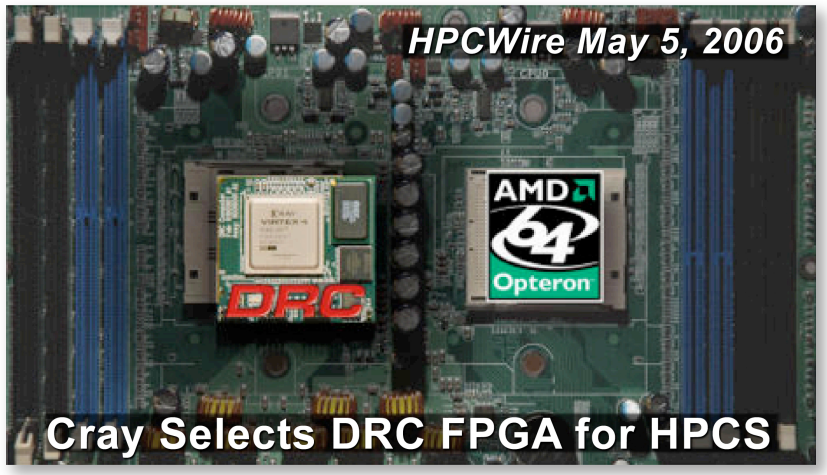
- Logic array: user-tailored to application
- On-chip RAM, multipliers & PowerPCs
- Gigabit transceivers/DSP blocks => FastIO/precision
- 100–1000 operations/clock cycle

Why FPGAs?

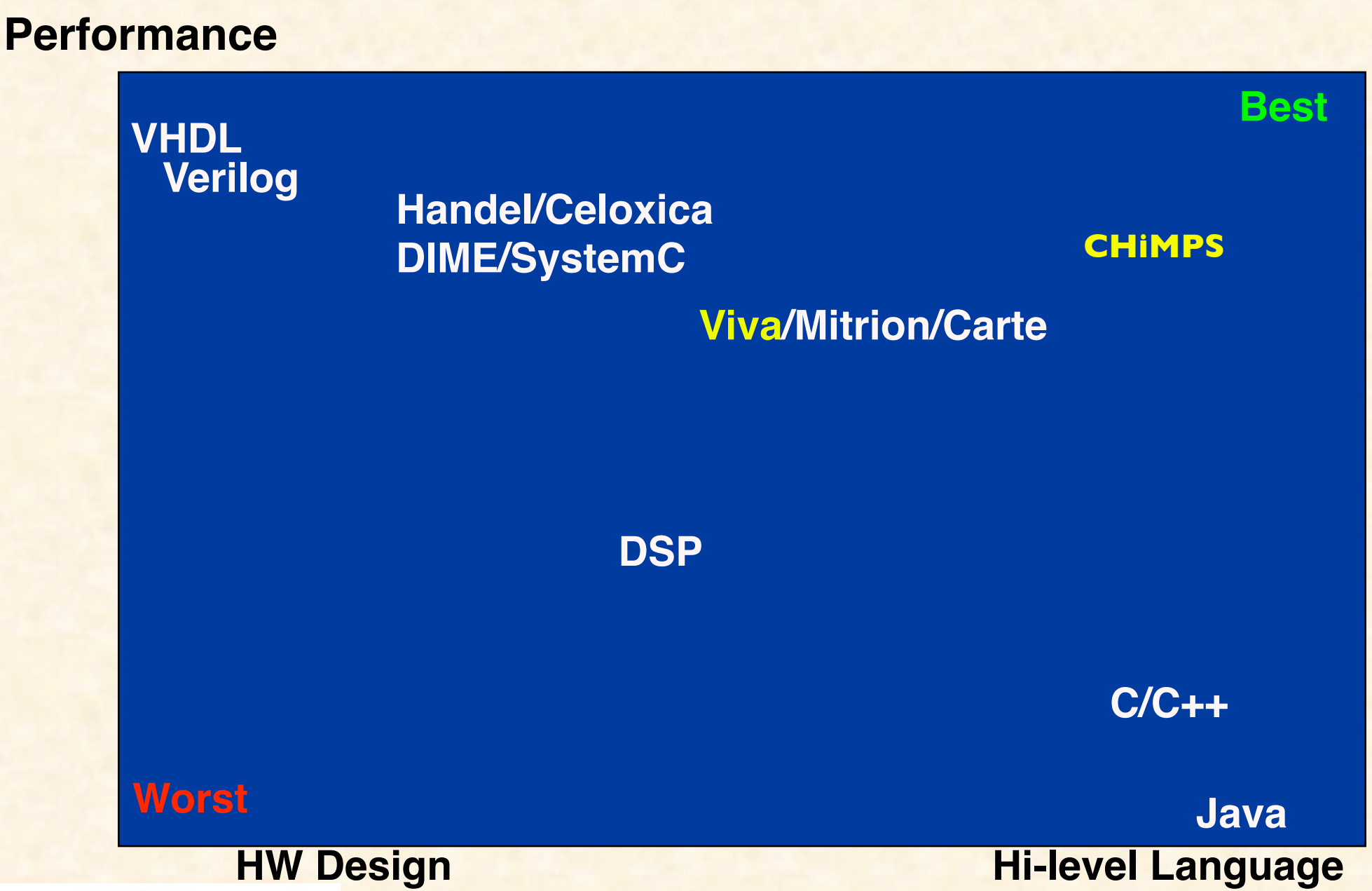
- **Performance**: optimal silicon use (maximize parallel ops/cycle)
- **Rapid growth**: Cells, Speed, I/O
- **Power**: 1/10th CPUs
- **Flexible**: *tailor* to application
- **Advances**: Telecom spinoff

High clock rate *is a cost, not a benefit; it drives up costs of everything else...*
-- eWeek



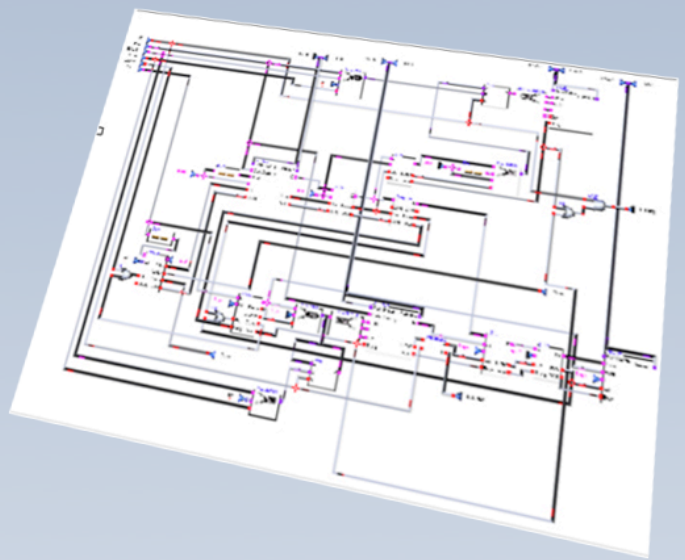


Tools: Performance vs. Coding Ease



Exploring programming options

Gauss matrix solver



Viva: Graphical Icons—3-dimensional

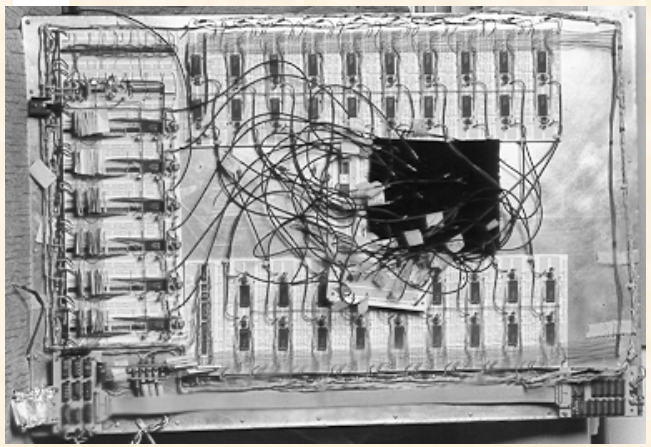
Compiler, simulator, and debugger

MitrionC: Text/flow—1-dimensional

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

Algorithms Developed

- **Matrix Algebra**: {V}, [M], $\{V\}^T\{V\}$, [M]x[M],GCD,…
- **n!** => Probability: Combinations/Permutations
- **Cordic** => Transcendentals: sin, log, exp, cosh…
- $\partial y/\partial x$ & $\int f(x)dx$ => Runge-Kutta: CFD, Newmark Beta: CSM
- **Matrix Equation Solvers**: [A]{x} = {b}, Gauss & Jacobi
- **Dynamic Analysis**: $[M]\{\ddot{u}\} + [C]\{u\} + [K]\{\dot{u}\} + NL = \{P(t)\}$
- **Nonlinear Analysis**: reduces NL time
- **Analog Computing**: digital accuracy
- **Structural Design/Optimization**
- **Unsolved App**: Traveling Salesman

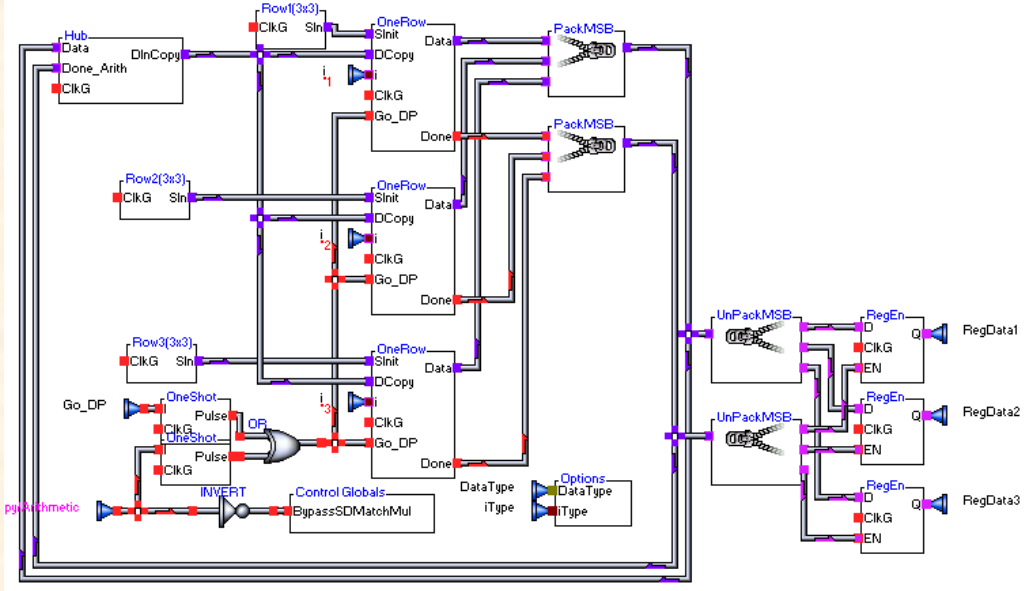


Future Technologies Group ornl

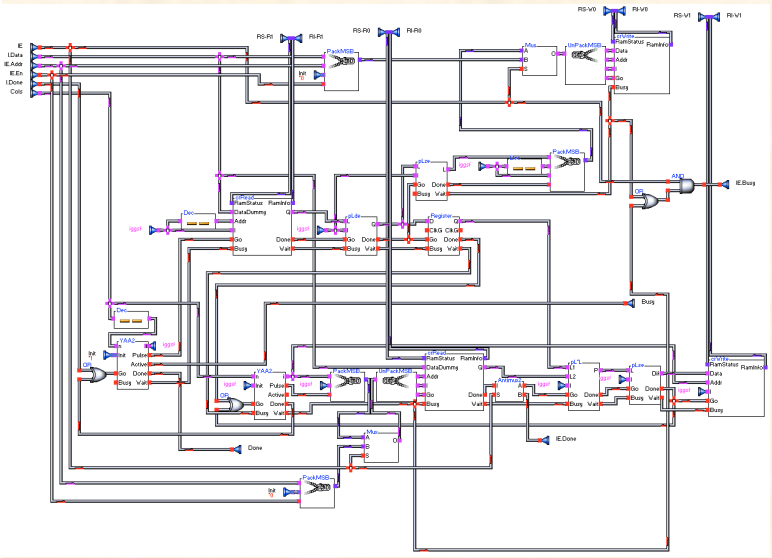
OAK RIDGE National Laboratory

Applications: VIVA Code

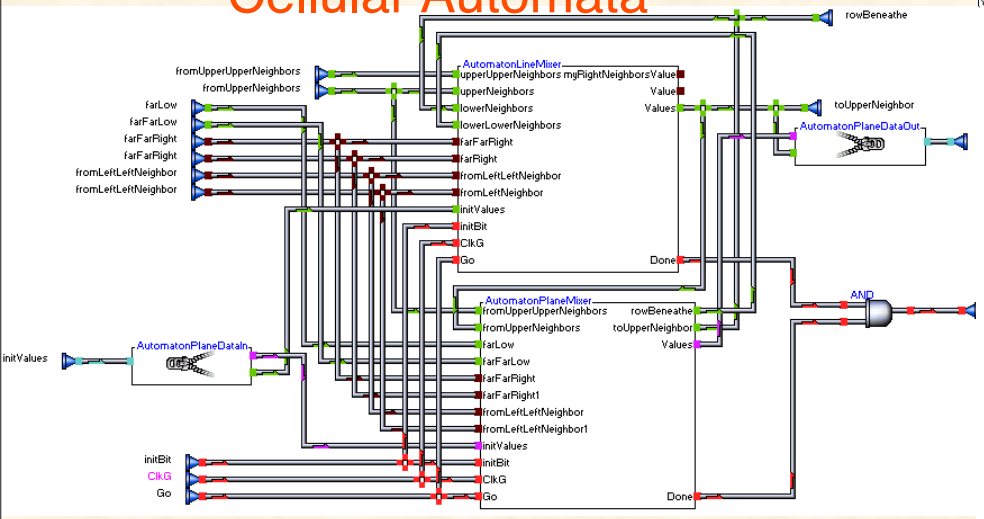
Jacobi Matrix Solver



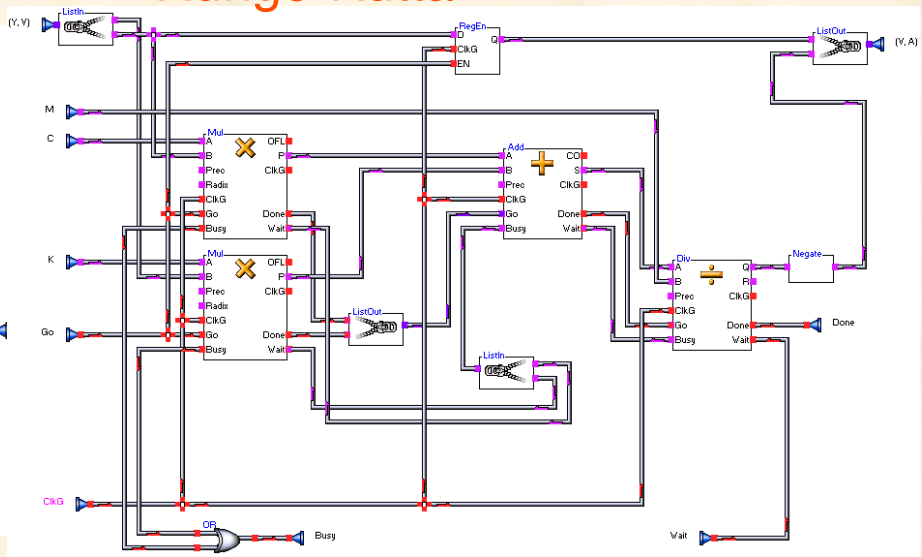
Gauss Matrix Solver



Cellular Automata



Runge-Kutta



Future Technologies Group ornl

OAK RIDGE National Laboratory

Applications

- Genomics
- Matrix Equation Solution
- Molecular Dynamics, Weather/Climate

Openfpga.org Smith-Waterman Benchmark

- ***FASTA*** ***(University of Virginia) application***

  http://fasta.bioch.virginia.edu

- **Uses search34 code & Cray SWA 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

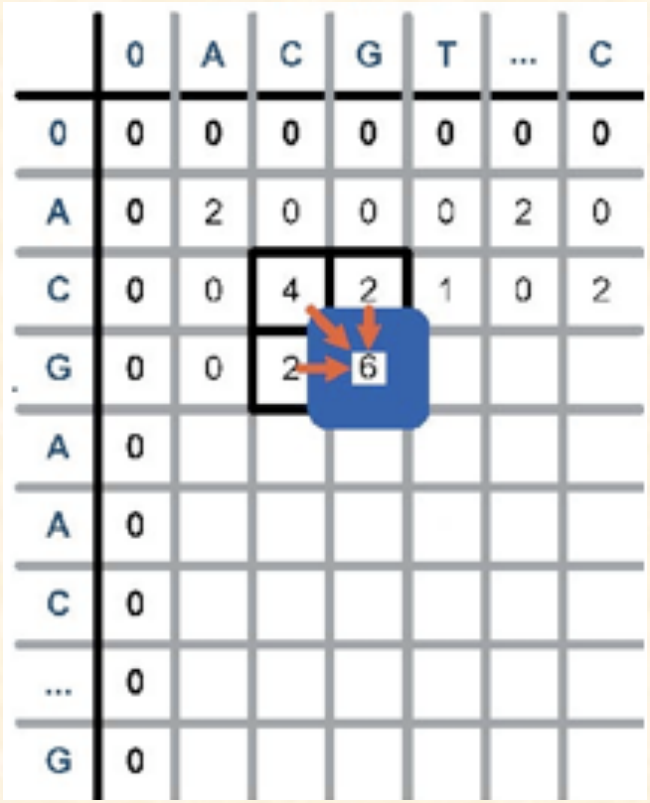
Final alignment

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

Smith-Waterman Algorithm Scoring

Query Sequence

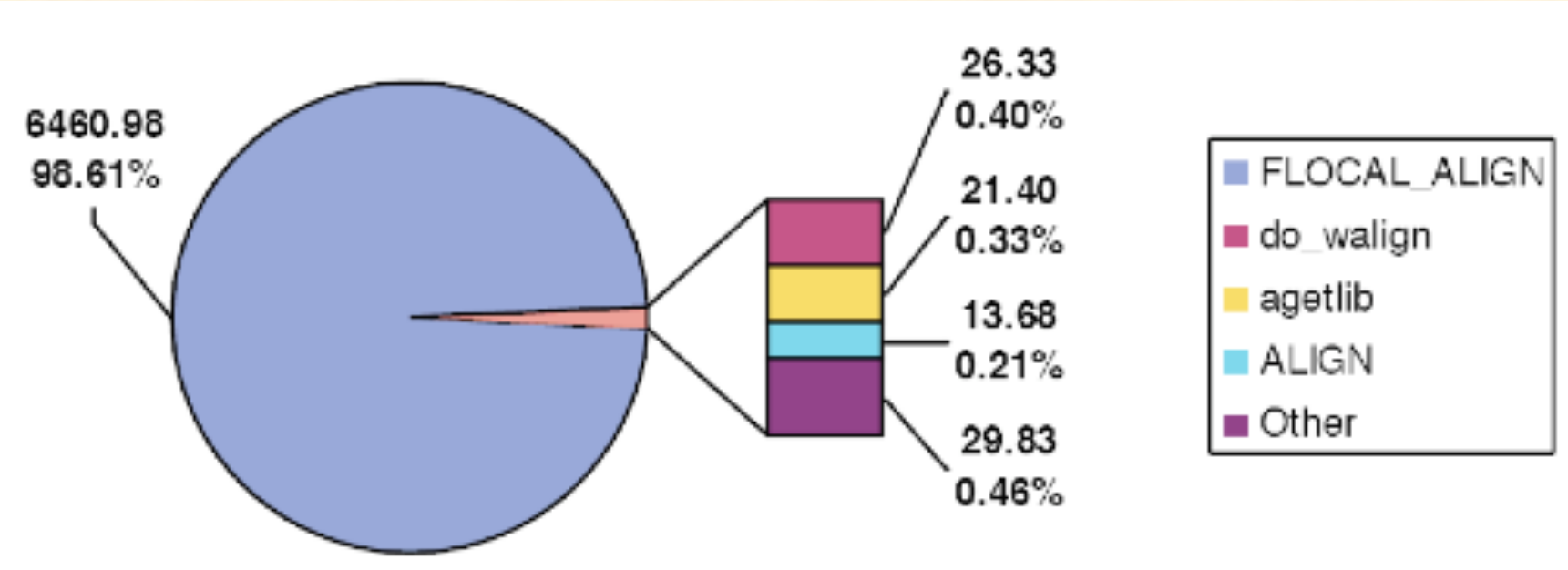
Database Sequence



	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						

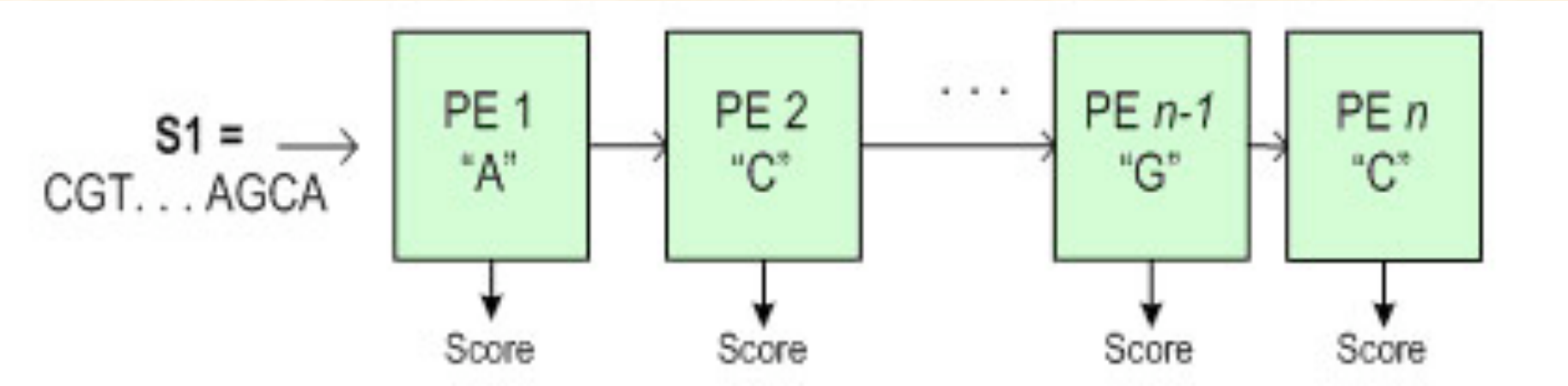
1. Initialize row & column 1 to 0
2. Score matches from upper left
3. Add to above-left score *(2+4=6)*

Search34 Computation Profile



98.61% is FLOCAL_ALIGN => VHDL kernel

Smith-Waterman Pipeline



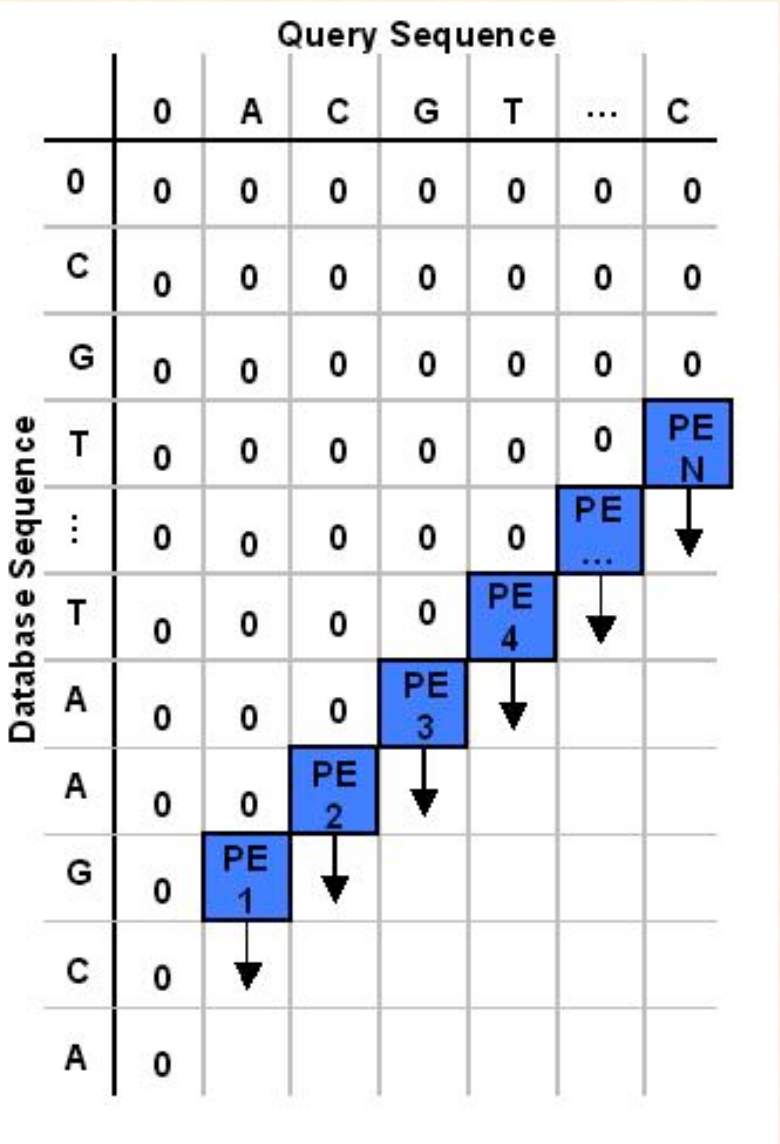
1. Query character preloaded into each PE
2. String S1 shifted thru pipe to compare
3. Score generated

Future Technologies Group ornl

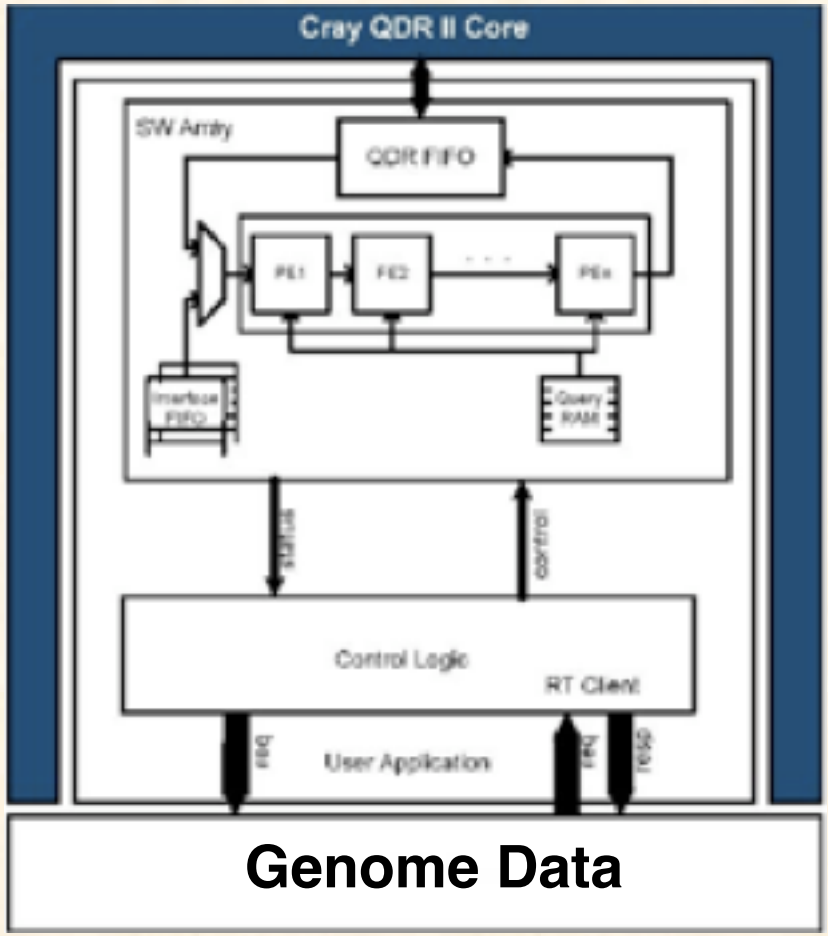
OAK RIDGE National Laboratory

Smith-Waterman

Parallel Score Calculation



Overall Algorithm



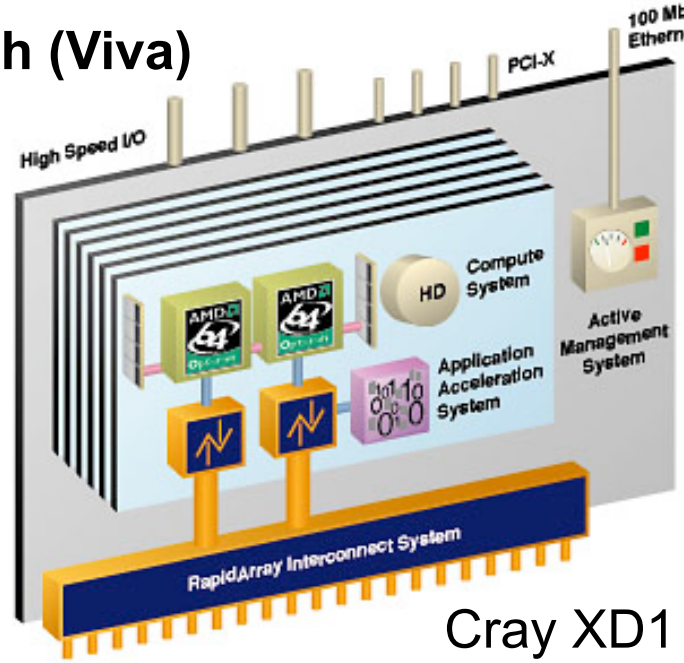
Future Technologies Group ornl

OAK RIDGE National Laboratory

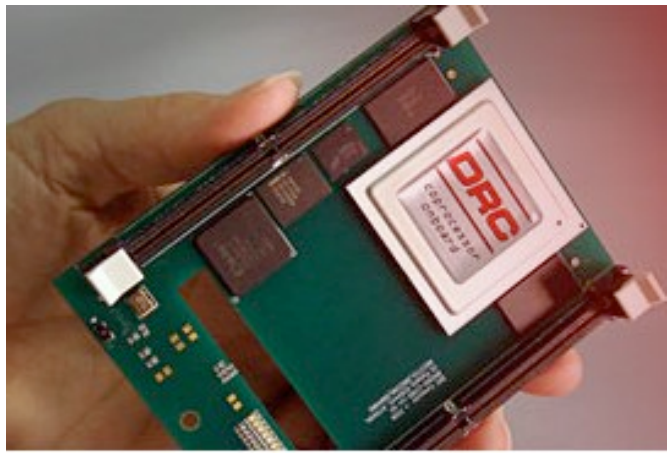

IBM Cell BladeCenter

ORNL FPGA hardware/tools

- SRC-6 (Carte), Digilent (Viva, VHDL), Nallatech (Viva)
- Cray XD1 (MitrionC, VHDL):
  6 FPGAs + 144 Opterons
- SGI RASC-Altix/Virtex4s (MitrionC)
- CHiMPS (Bee2 => Cray XD1 => DRC => XT4)
  (Xilinx early access)



Cray XD1



FPGA Performance

ORNL XD1 (Virtex2): Initial Results

Case 1: *Micro-RNA*

FPGA vs Opteron Time (hrs) for FASTA

	1	2	3	4	5
CPU 2.2GHz	75	-	-	-	-
FPGA(s) 0.2GHz	7.39	3.75	2.48	1.91	1.56
FPGA Speedup vs 1 CPU	10.15	20.0	30.2	39.3	48.1

Output Options (Impact Speedup)

Detailed: -Q -H -f -10 -g -3 -d 10 -b 10 -s OpenFPGA.mat -E 0.0001
Minimal: -Q -H -f -10 -g -3 -d 0 -b 10 -s OpenFPGA.mat -E 0.0001

Cray XD1 FPGA Speedup vs. 2.2 GHz Opteron

Case 2: *Bacillus anthracis DNA comparison*

Virtex2 Pro 50 Speedup

	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	Avg	SD
8k	26	30	27	32	30	30	29	30	30	27	31	29	31	30	31	31	31	30	29.6	1.2
16k	22	25	26	31	30	30	28	31	28	27	30	29	29	29	32	31	32	29	28.7	2.5
8k	49	49	49	50	49	49	50	49	49	49	49	49	49	49	50	49	49	49	49.4	0.2
16k	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	49	50	50	49.9	0.3

Virtex4 LX160 Speedup: 8 hrs => 5 mins*

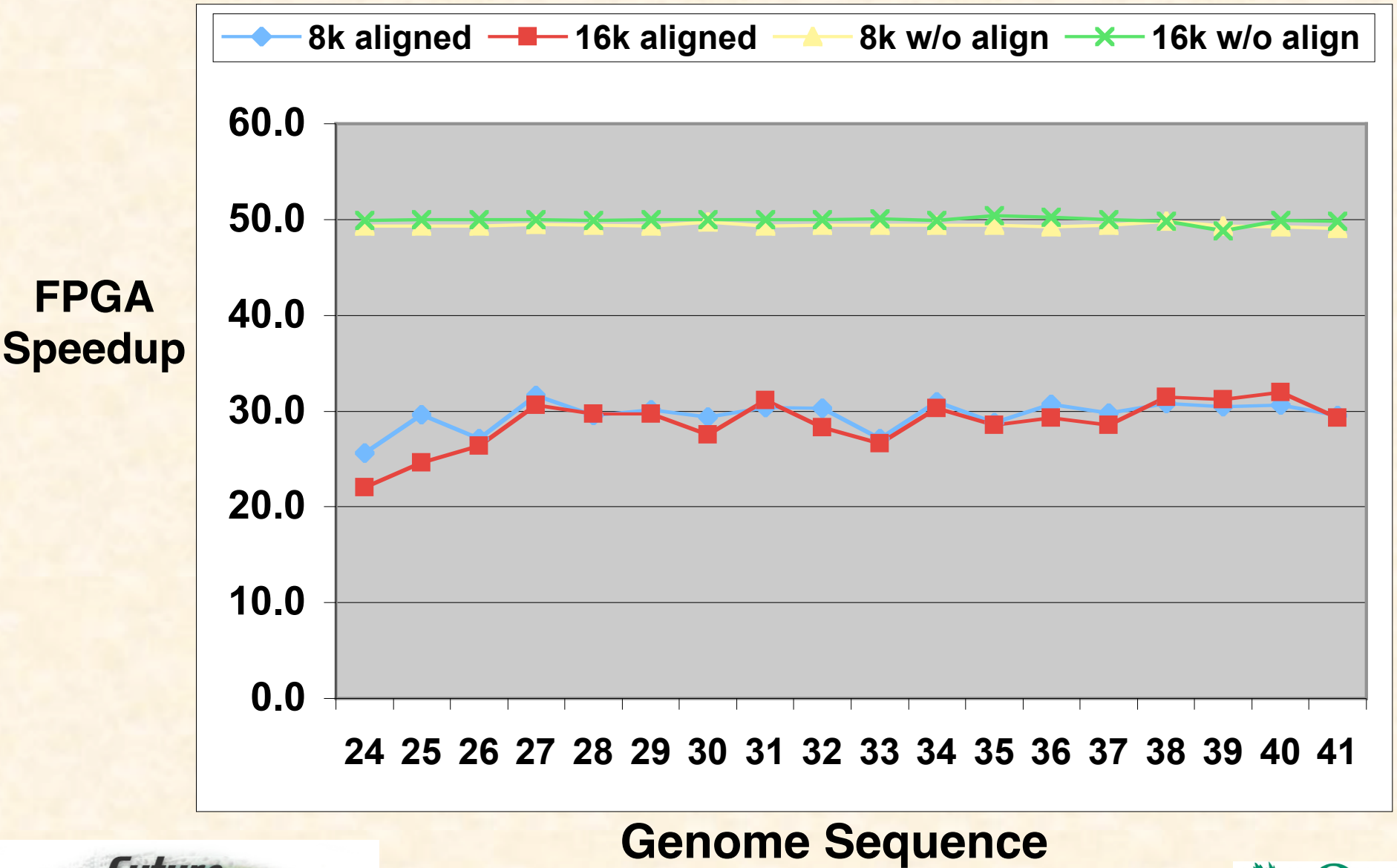
	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	Avg	SD
8k	36	43	39	47	44	45	43	45	45	39	46	42	46	44	46	46	46	43	43.5	2.9
16k	29	33	37	45	44	43	39	47	41	37	46	41	43	41	47	46	48	43	41.5	4.9
8k	98	98	98	97	97	98	98	98	98	97	98	98	98	98	98	98	98	97	97.6	0.1
16k	100	101	101	100	100	100	101	101	101	101	101	101	101	101	100	100	101	100	100.7	0.4

*28,873 => 288 secs

24 => Sequence AE017024

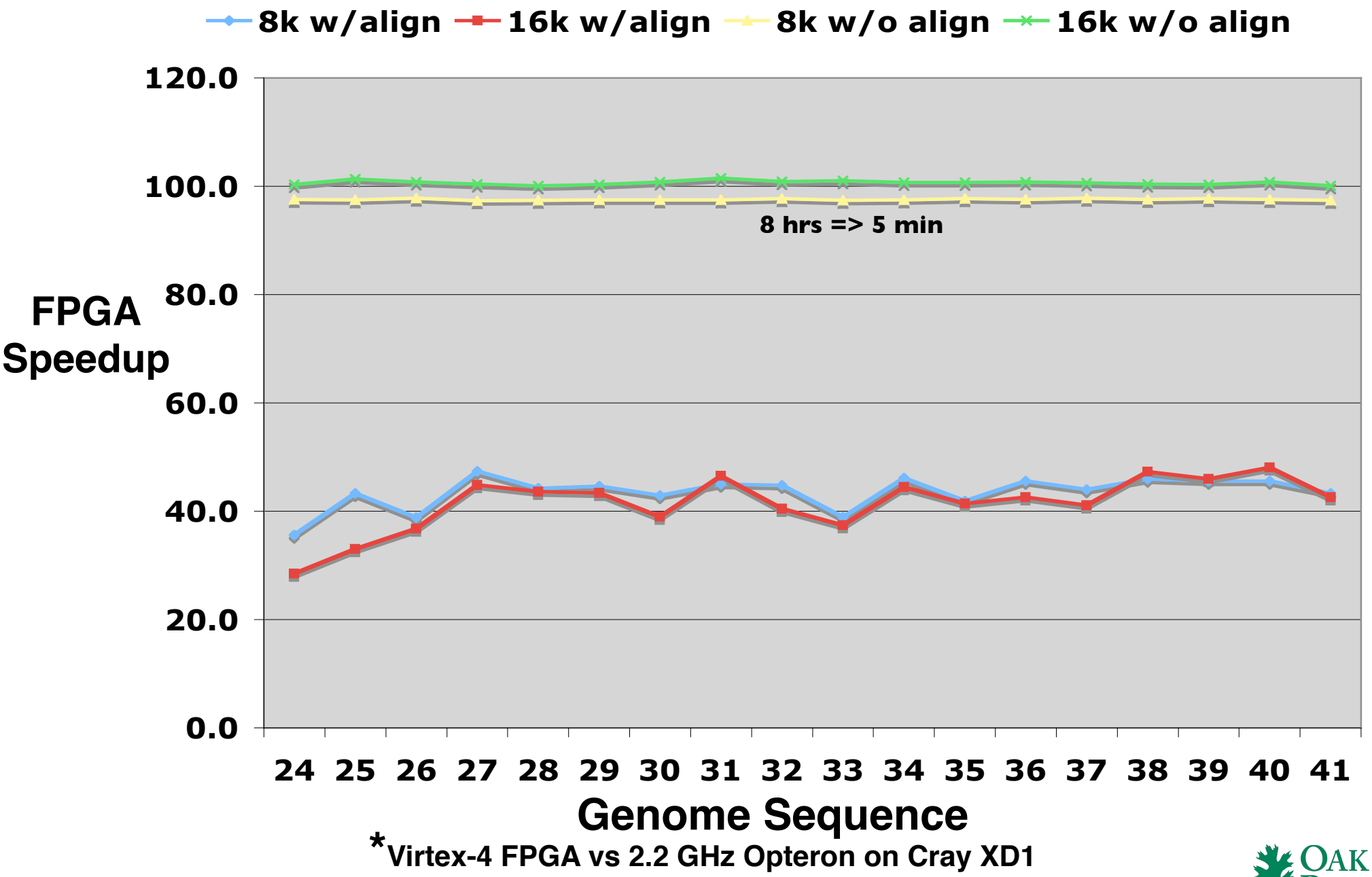
XD1 Virtex2 Speedup vs. 2.2 GHz Opteron

Case 2: *Bacillus anthracis DNA comparison*



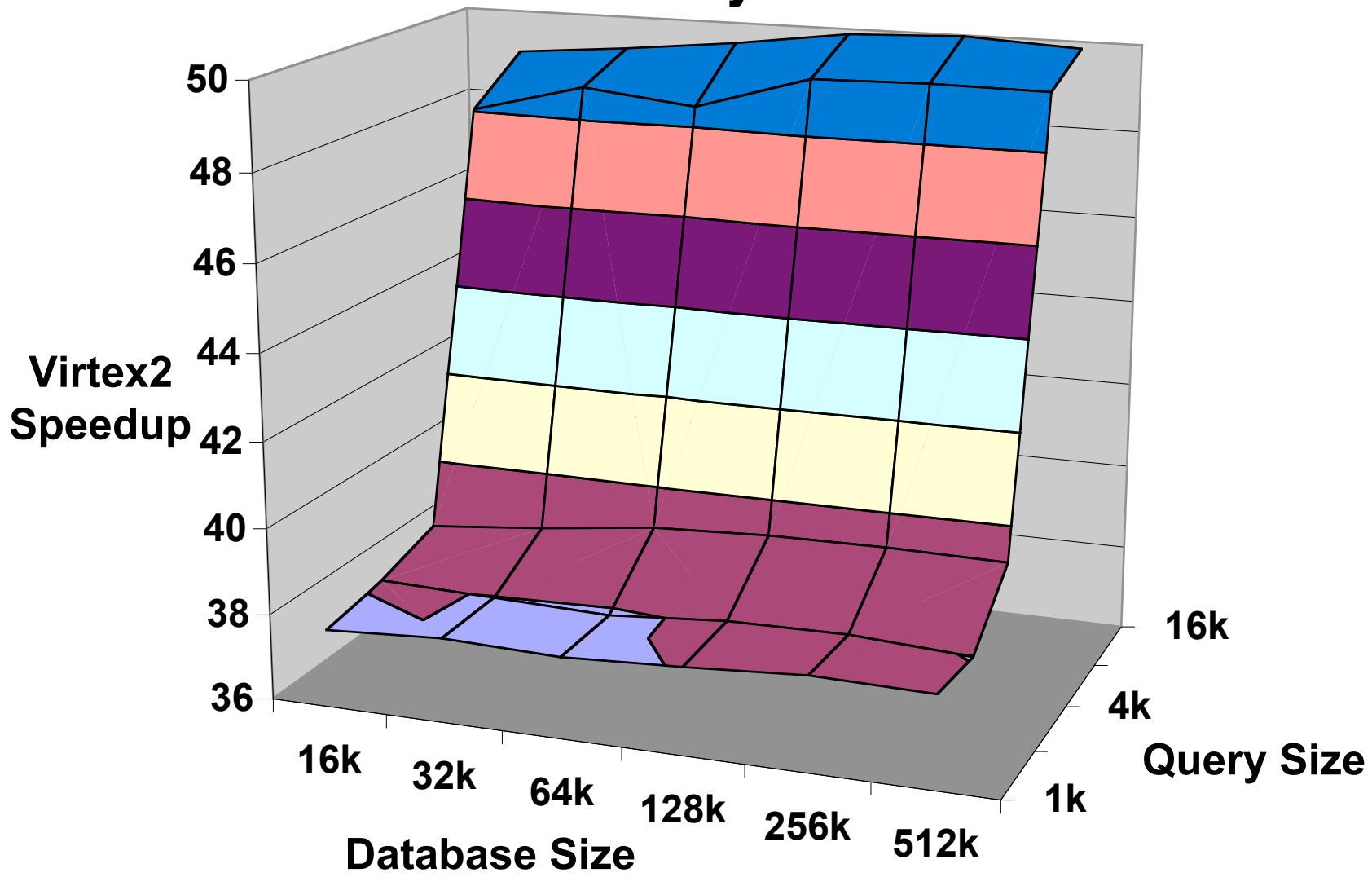
100x* DNA Sequence Speedup

Bacillus anthracis Human DNA comparison



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

FPGA Speedup Grows with Query Size



1st Results for 150 FPGAs*

Solved huge DNA sequencing problem:

12.5 years (150 mos.) for 1 Opteron

6 weeks for 150 Opterons

1 day for 150 FPGAs

7,350X Speedup over one 2.2 GHz Opteron

Next: Test performance on *64 Virtex4 FPGA Maxwell* at Edinburgh University

*Thanks to NRL for use of 150 FPGA Cray XD1

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

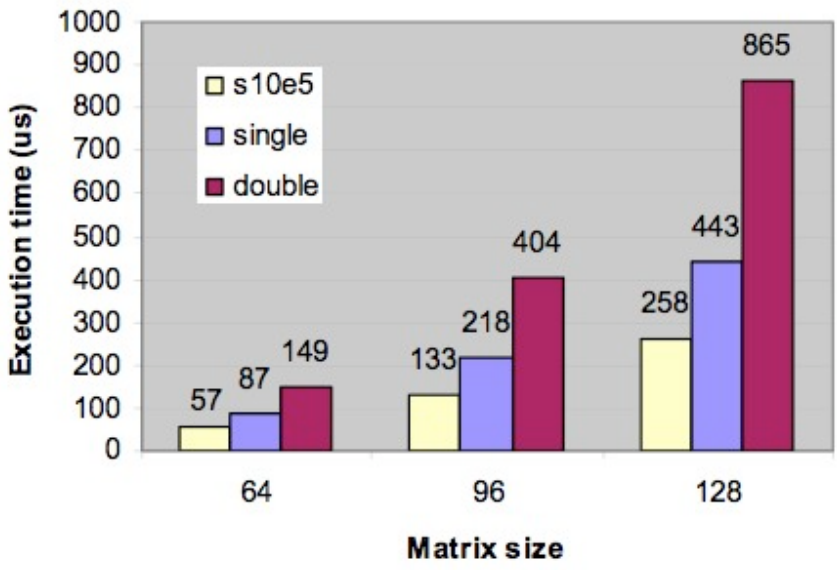
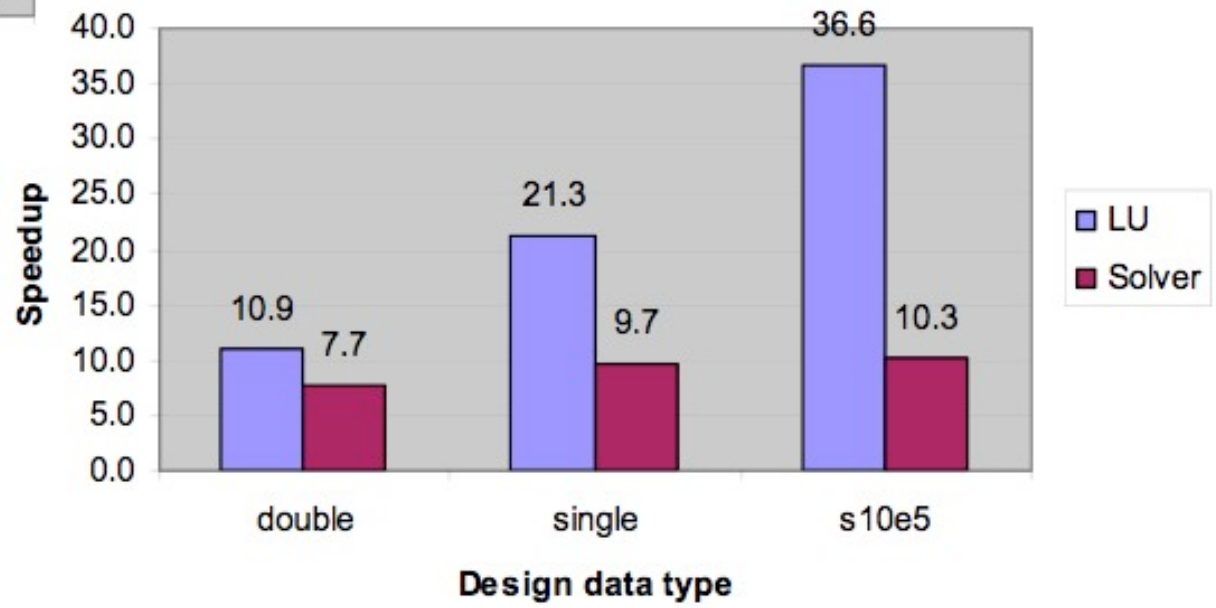


Table 6: LU implementation on XC2VP50-7

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%)



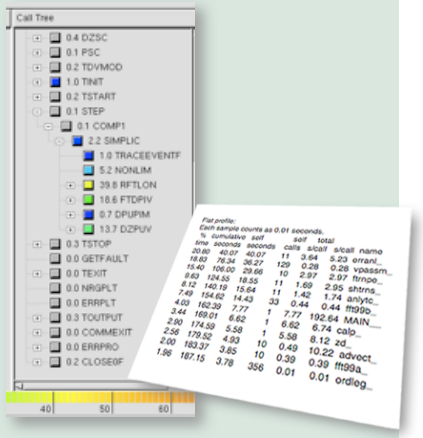
Benefits:
High performance **of LP arithmetic**
High precision **accuracy**
Speedup increases **with matrix size**
(LU dominates calculations)

First mixed-precision LU & solver for FPGAs

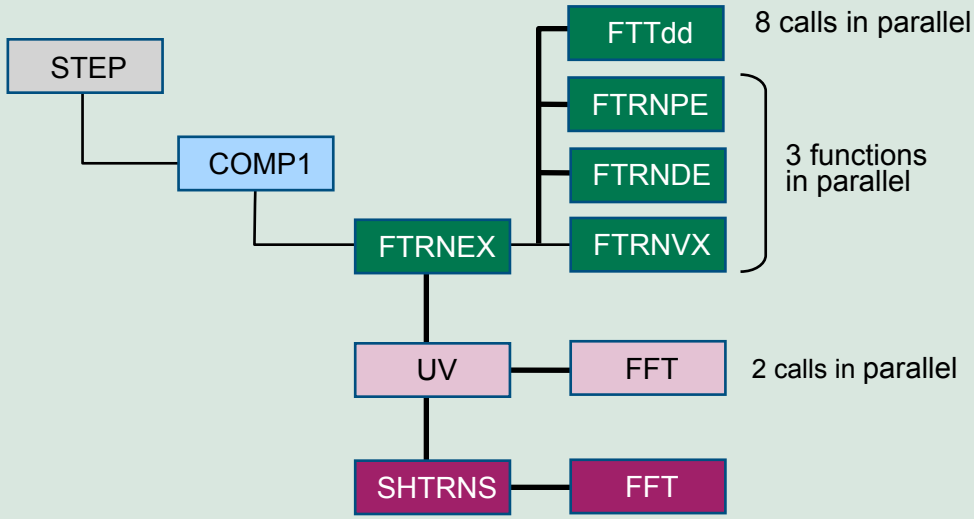
***Virtex-II vs 2.2 GHz Opteron**

Ported Weather-Climate code Spectral Transform Shallow Water Model (STSWM) to FPGAs

Profile

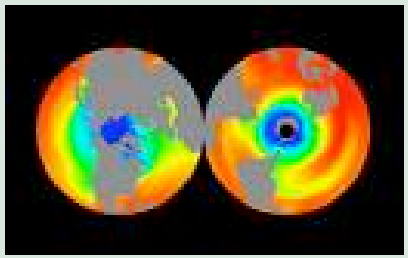


Find parallelism: 80% FFTs



Goal

More GF/$ GF/Watt



Model 5-10X faster

Summary

- **ORNL HPC & FPGA research:**
  - **FPGA growing in HPC**
  - **Partners:** *Cray, Xilinx, UT, Mitrion, NRL, SGI*
  - **FPGA Systems: Cray, SRC, Nallatech, Digilent, SGI**
  - **Compilers: Mitrion-C, Carte, Viva, DSPlogic, CHiMPS**
  - **Speedup: 10X Eqn Soln, 100X DNA Sequencing**
  - **Scalable: 7350X Speedup for 150 FPGAs (Genomics)**
- **Next: Maxwell, CHiMPS & DRC (Cray Module)**

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

Contact

Olaf O. Storaasli
Future Technologies Group
Google **Olaf ORNL**

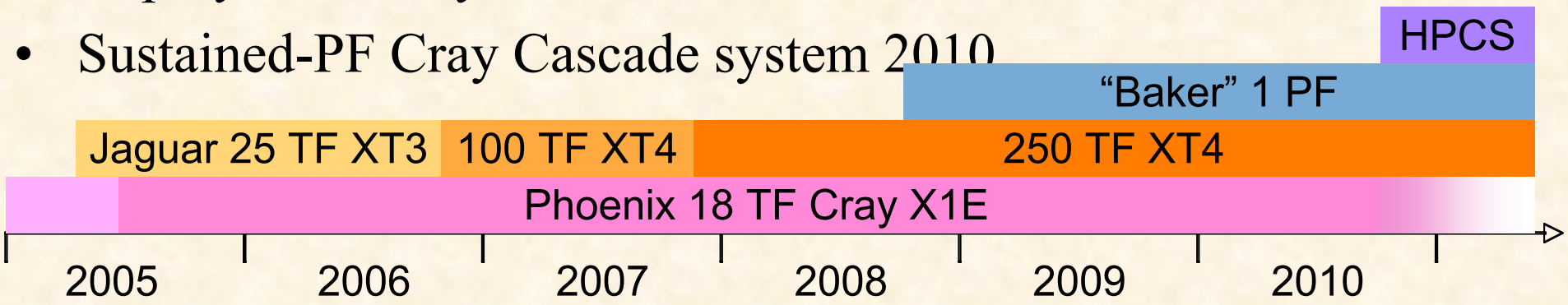
THANK YOU

OAK RIDGE National Laboratory

ORNL Milestones: Deliver 1PF system in 2008 Deliver 250 TF by 2007

Roadmap

- Upgrade existing 25 TF XT3 to dual-core 100 TF system in 2006
- Upgrade 100 TF to 250 TF in late-2007
- Deploy 1 PF Cray "Baker" late 2008
- Sustained-PF Cray Cascade system 2010



Results: 150 2.2GHz Opterons @ NRL

Time to Completion

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

Consistent with CUG07 results