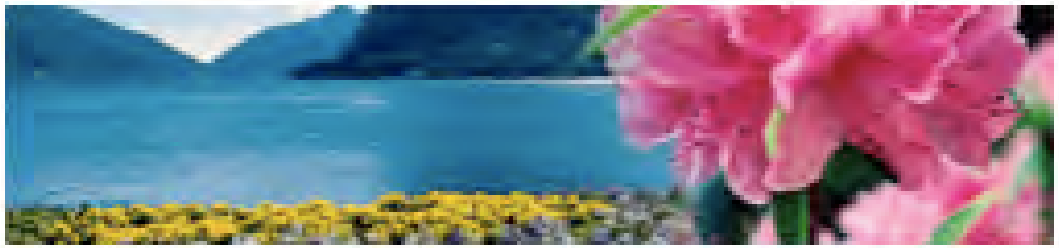
Experiences Harnessing Cray XD1 FPGAs

& comparison with other FPGA HPC Systems



Olaf O. Storaasli, Melissa S. Smith & Sadaf Alam
Oak Ridge National Laboratory, Tennessee

Future Technologies Group ornl

OAK RIDGE National Laboratory

Status Report

Are FPGAs ready for HPC *Prime-Time*

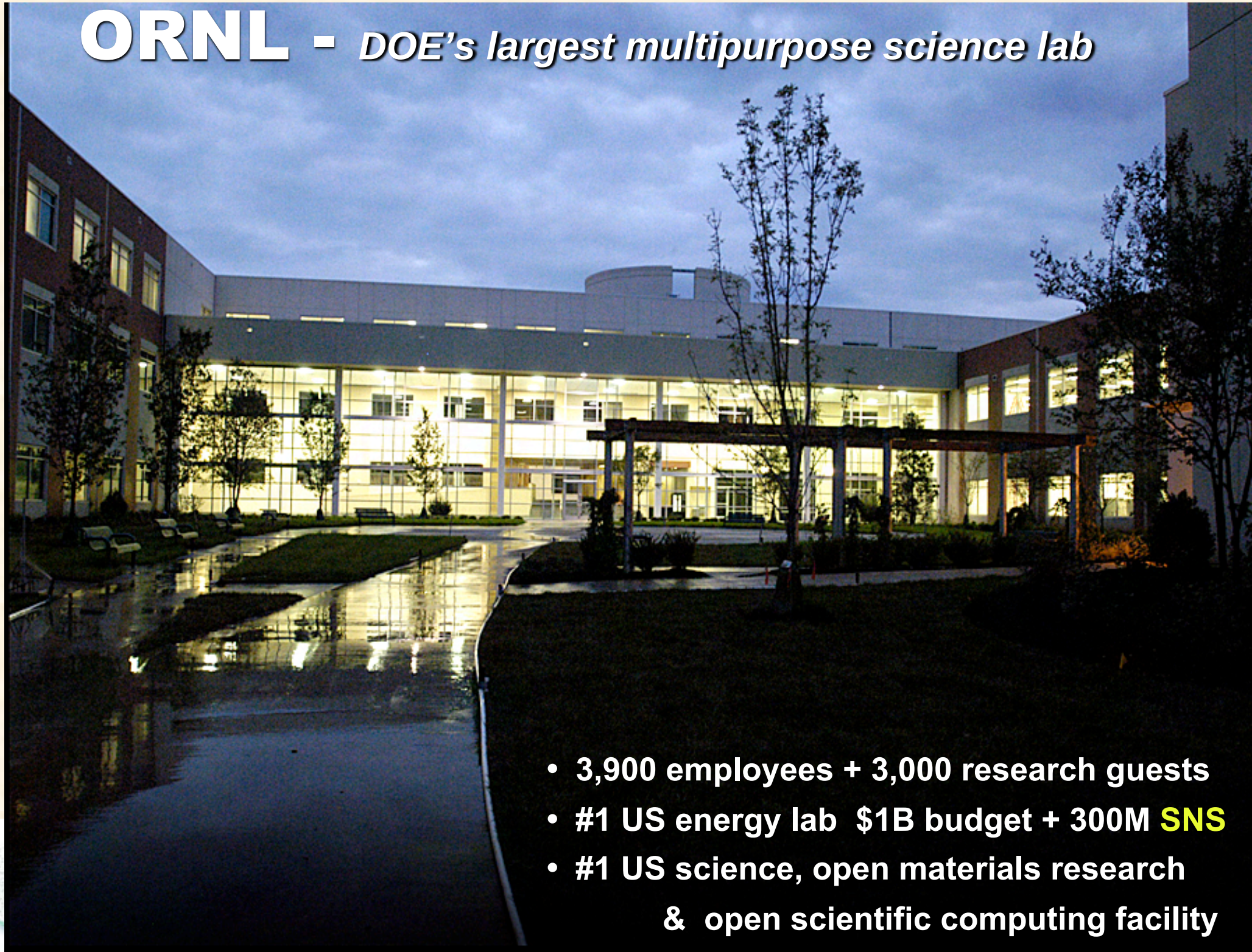
Outline

- **Background: ORNL, FT, FPGAs, SW, Apps**
- **Cray XD1 HW & SW**
- **VHDL vs HLL**
- **Mitrion-C tests: Virtual & Real**
- **Comparisons: SRC (Carte), SBS (Viva)**
- **Summary**

Future Technologies Group ornl

OAK RIDGE National Laboratory

ORNL - *DOE's largest multipurpose science lab*



- 3,900 employees + 3,000 research guests
- #1 US energy lab $1B budget + 300M SNS
- #1 US science, open materials research & open scientific computing facility

ORNL: *21st century research campus*

ORNL High-Performance Computing

Computer Room 180˚ Panorama **Cray XD1**

OAK RIDGE CENTER FOR ADVANCED STUDIES

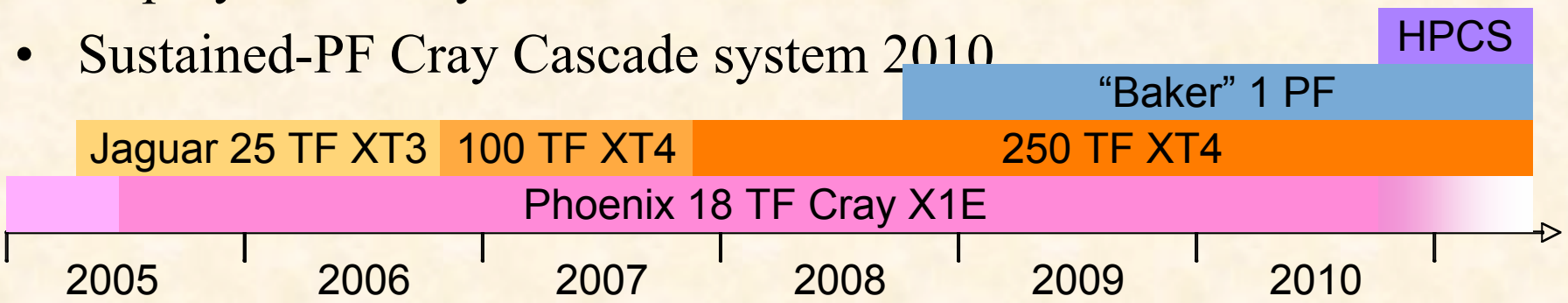
Future Technologies Group ornl

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



- 18 TF Cray Phoenix and 25 TF Cray Jaguar currently in production

Future Technologies Group ornl

OAK RIDGE National Laboratory

Future High-Performance Computing

Volatile - novel emerging architectures

Traditional

Commodity: 2^n core chip clusters, ~2 GHz

Special: MTA, El Dorado, Cyclops, PIM

To watch

- **FPGA**: DSP => HPEC => HPC <==
- **Cell**: IBM, Sony, Toshiba
- **Array**: Clearspeed
- **Graphical**: Nvidia, ATI 100x speedup
- **Optical**: Lenslet 56,624x speedup

Future Technologies Group ornl

OAK RIDGE National Laboratory

Cray CTO view of HPC Future:

- *After exhaustive analysis, Cray Inc. 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."* (HPCWire, 3/24/06).

The Register®

Software | Hardware | Internet | Telecoms | Mobile | Security | Management

Servers | Storage | Data Networking | Reg Reader Studies | Virtualisation

Cray places order for Opteron Helper

FPGA supercomputer booster

By Ashlee Vance in Austin

Published Wednesday 3rd May 2006 18:08 GMT

Cray has turned to a small start-up for possible help with its grand plan to w
contract from the US government.

Cray this week announced that it will use a specialized coprocessor from D
Opteron-based systems. The DRC product – an FPGA (field programmable
will allow Cray's supercomputers to churn through specific workloads such
subroutines at a much quicker pace than standard Opteron chips alone. Cra
ready servers as one piece of its "Cascade" project that is competing with ri
Microsystems and IBM for a piece of a $200m DARPA contract.

Breaking News:

Cray Selects DRC's FPGA Coprocessors for Supercomputers

DRC Computer Corporation has been selected by Cray to provide a new coprocessor module as a massively parallel reconfigurable option for future Cray supercomputers.

DRC makes a coprocessor module that plugs into a standard AMD Opteron sockets, providing direct access to adjacent DDR memory and Opteron processors at HyperTransport speed and nanosecond latency. Tight coupling between CPU and memory means that bandwidth and latency bottlenecks are virtually eliminated, allowing 10x to 100x performance improvement while lowering power and heat requirements. These benefits apply to the full range of high-performance computing systems, from entry level to the world's largest supercomputers.

Performance: FPGAs vs. CPUs

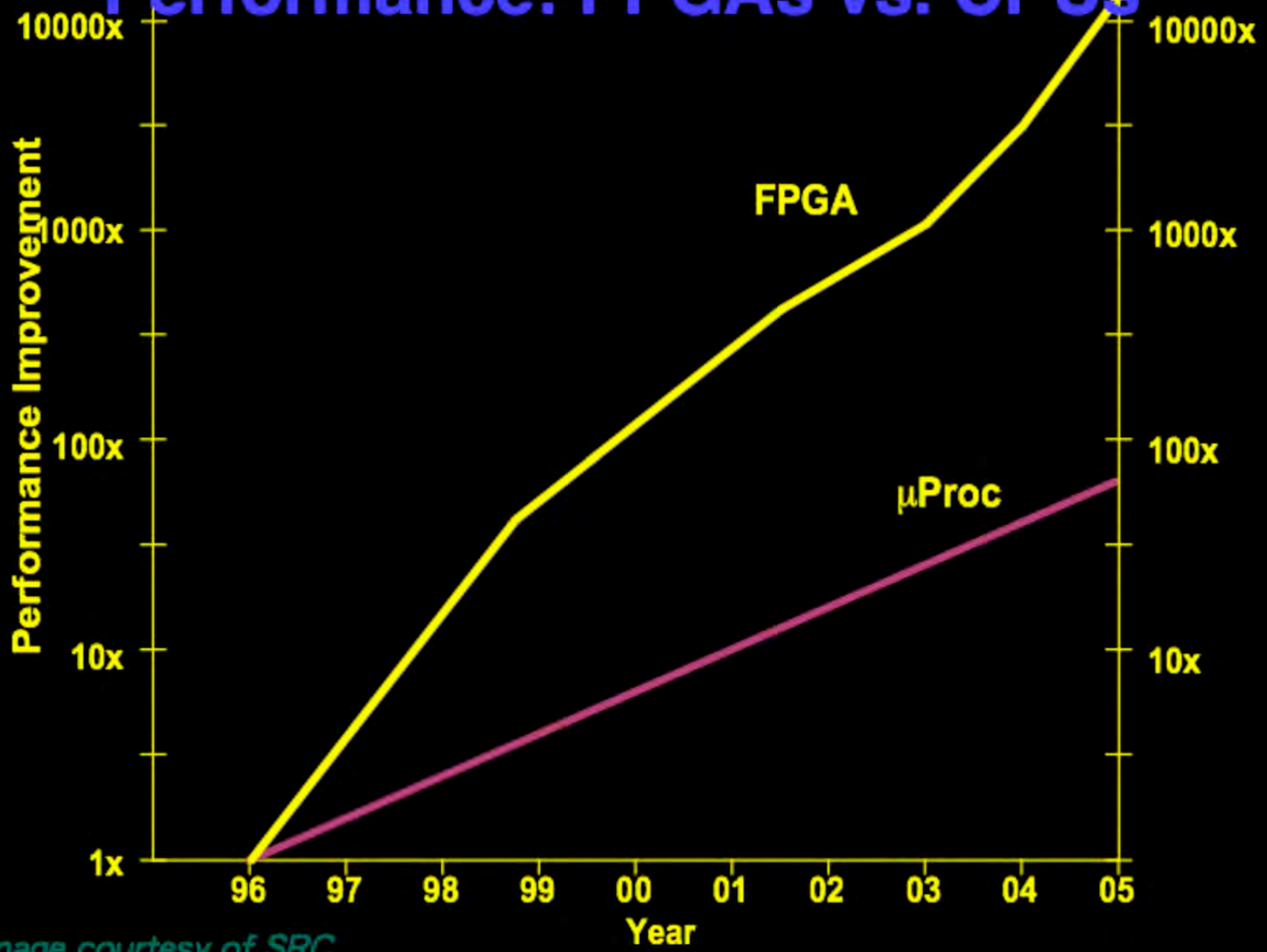


Image courtesy of SRC

ORNL Experimental Computing Laboratory (ExCL)

Managed by the Future Technologies Group

FPGA-based Systems

- **Cray XD1**: 144 2.2GHz Opterons + 6 Virtex-II Pro FPGAs connected via HyperTransport.
- **SRC-6C**: dual 2.8GHz Xeons + Xilinx Virtex-II FPGA connected via DIMM slots.
- 3 **Digilent Virtex-II Pro FPGA** Systems with Viva
- **Nallatech XtremeDSP**: Virtex-II Pro FPGA with Viva.

Other Novel Systems

- **Cell**: dual 2.4GHz Cell processors, with 64-bit Power Architecture PPE core & 8 SPE SIMD cores.
- **GPU accelerators**: NVIDIA NV3x/2 Opterons & NVIDIA NV4x/2 Xeons.
- **OpenSSI cluster:** 31 2.6GHz Xeons, 10/100 Ethernet, as Lustre and OpenSSI testbed.
- 2 **ClearSpeed CS-301 PCI boards**: 2 Array Processors with 64 parallel execution units.
- **Iwill H8501 server:** 8 1.8GHz dual-core Opterons, 32 GB memory on NUMA HT - 16-way SMP.
- **4 compute servers**: each with 2 dual-core Opterons, & dual processor, 500 GB RAID fileserver.

ORNL Cray XD1

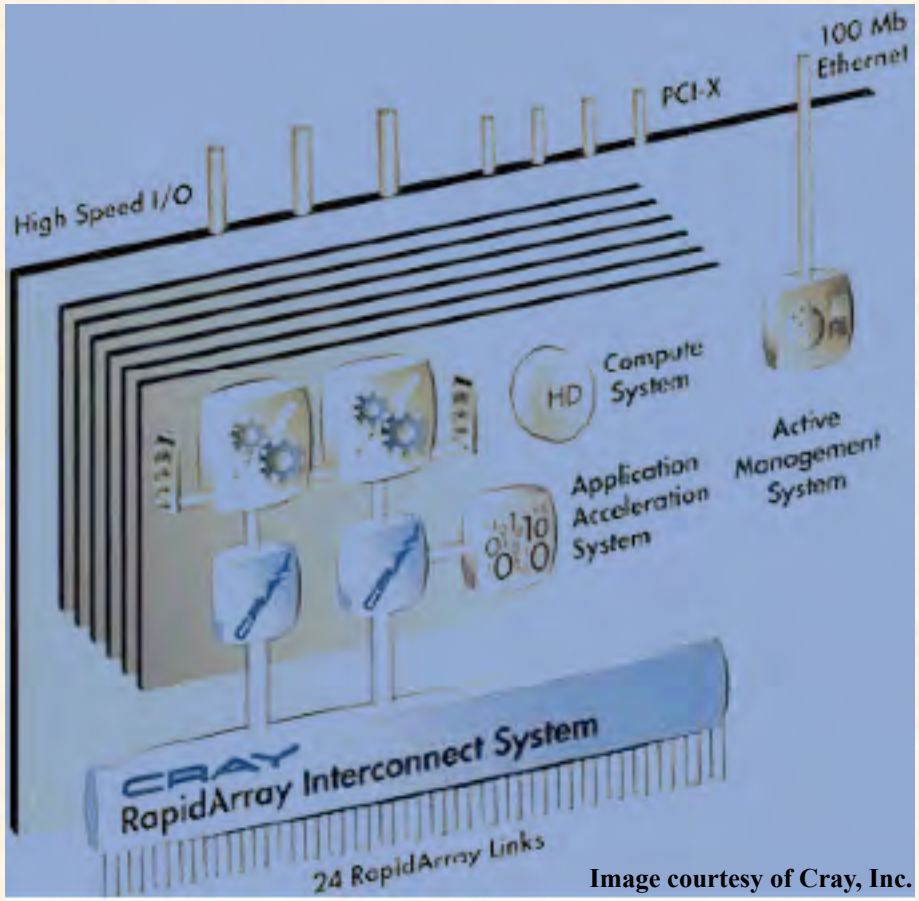


Image courtesy of Cray, Inc.

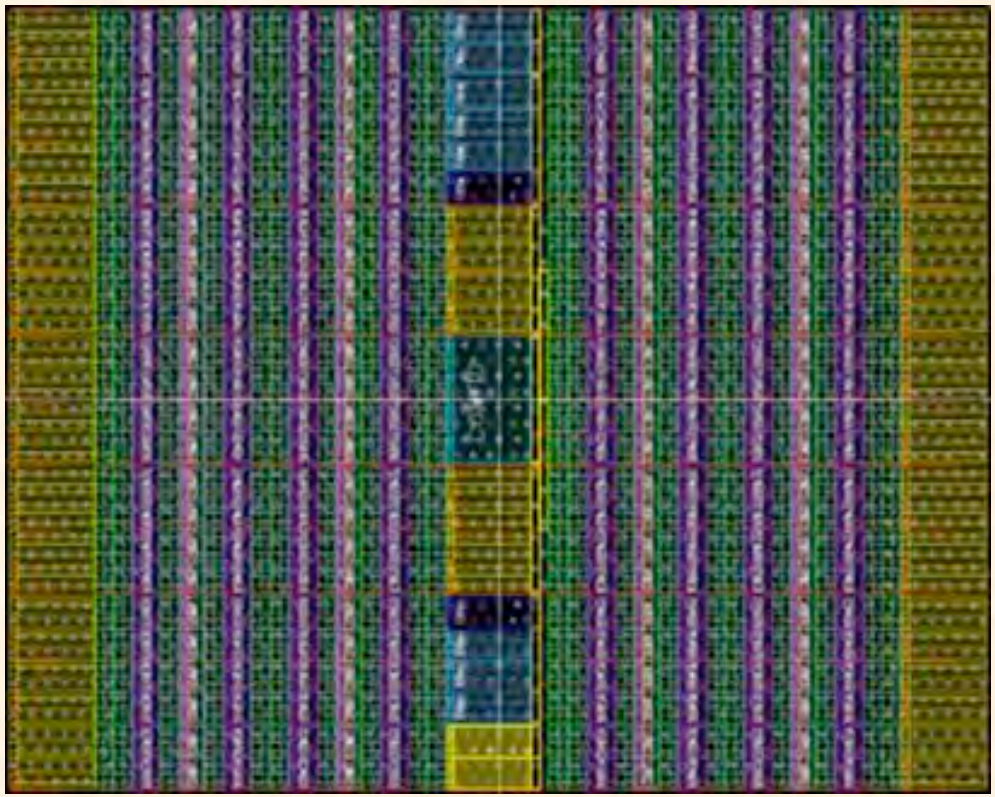
- Installed Oct '04
- 12 Chassis with 144 processors
  (2.2GHz AMD Opteron 248 processors)
  => 633 GFLOPS peak
- 576 GB Memory, 18TB disk
- 6 XCVP50 Xilinx Virtex2Pro FPGAs (5M gates each) in lower chassis
- Cray RapidArray terabit backplane
- Mitrion-C received Oct '05

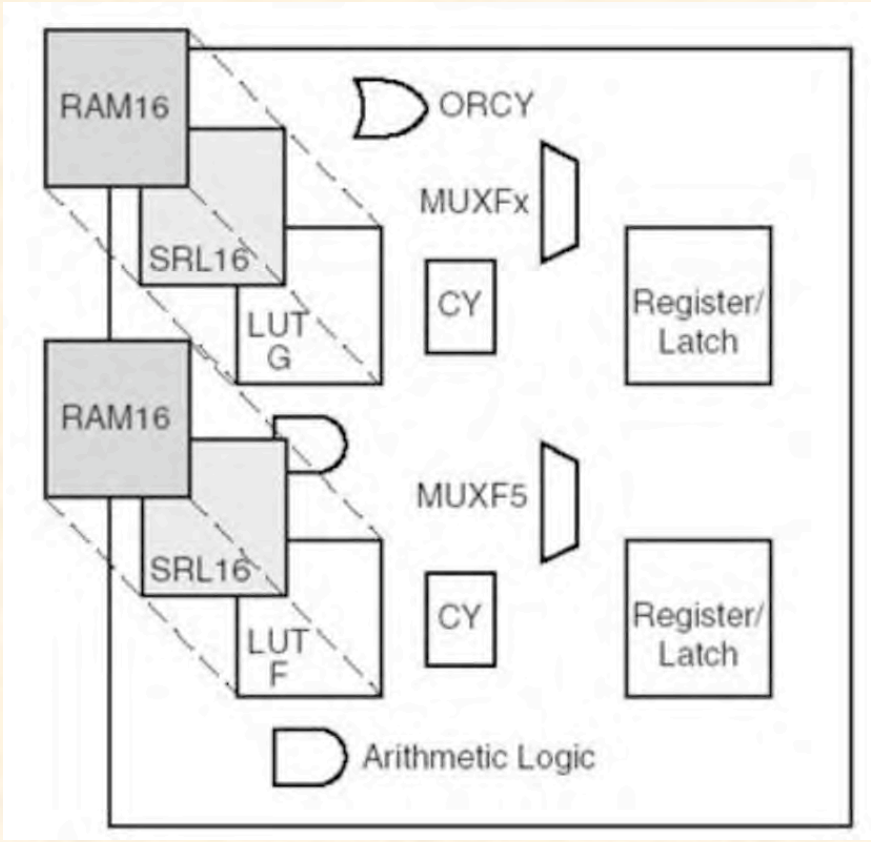
- Apps planned: matrix equation solver (GPS), biology (AMBER), mechanics (NIKE3D)

Future Technologies Group ornl

OAK RIDGE National Laboratory

Field-Programmable Gate Array (FPGA)





Xilinx Virtex4 FPGA (25K slices) **FPGA Logic slice**

Future Technologies Group ornl

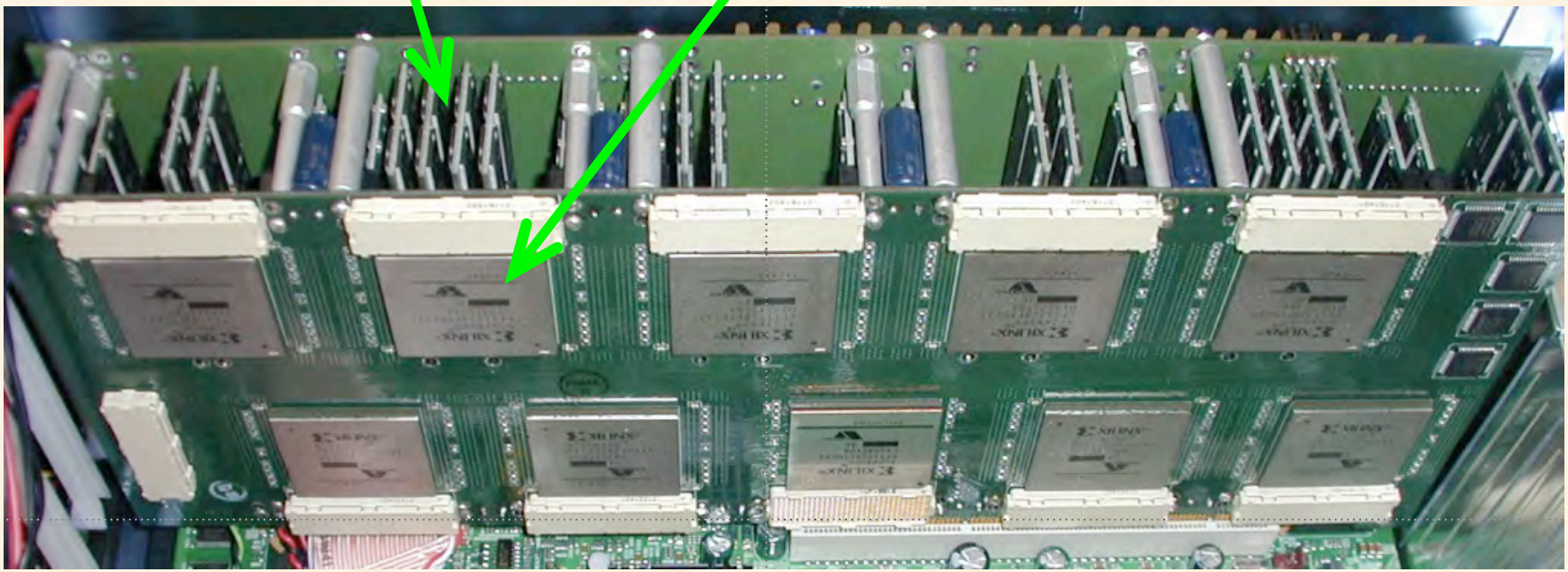
OAK RIDGE National Laboratory

Memory: FPGA & SDRAM

- keep "action" on/near FPGA -

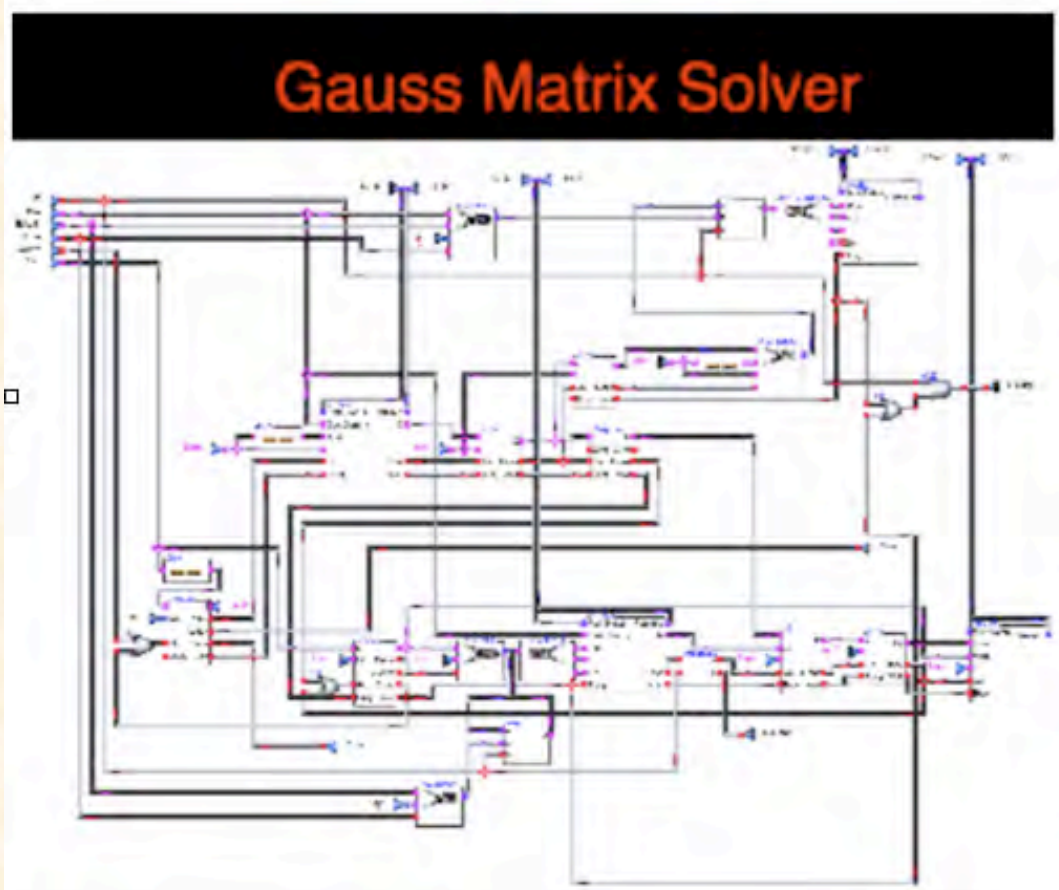
2-8GB SDRAM
(large applications)

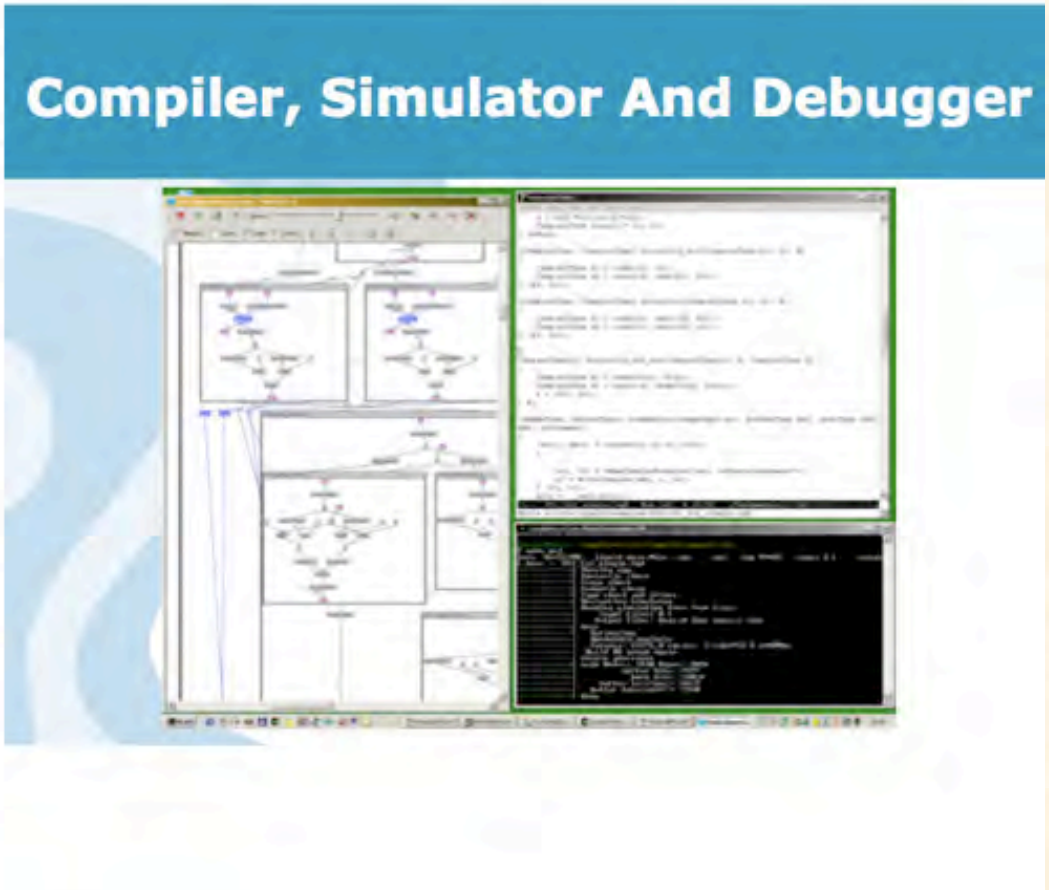
144x 2KB blocks RAM



- User configures in VIVA for any data type (fixed, float …)

Viva & Mitrion-C FPGA Code





Viva: Graphical Icons - 3D
(System Description)

Mitrion: PDK Text - 1D
(Virtual Processor)

Future Technologies Group ornl

OAK RIDGE National Laboratory

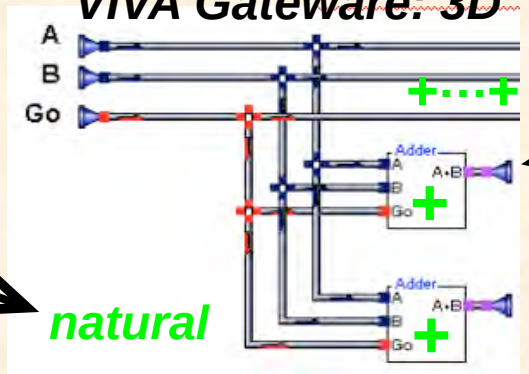
VIVA: *Custom Chip Design*

What: Graphically code FPGAs: *drag & drop* vs *text*)

Traditional Code: 1D

```
do i = 1, 1000
   C= A+B
end do
```

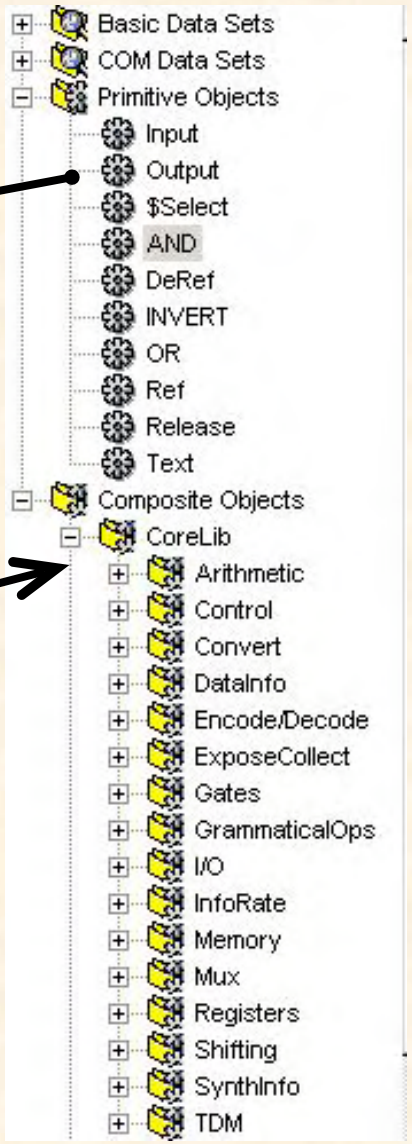
VIVA Gateware: 3D



Parallelism

esoteric ***natural***

VIVA Menu



How: Converts *icons-transports* to FPGA circuit

Why: near-ASIC speed (w/o chip design $$$)

Corelib: Pre-built objects & examples

Data: Any type-size-precision (not fixed)

System Description: ports to any hardware

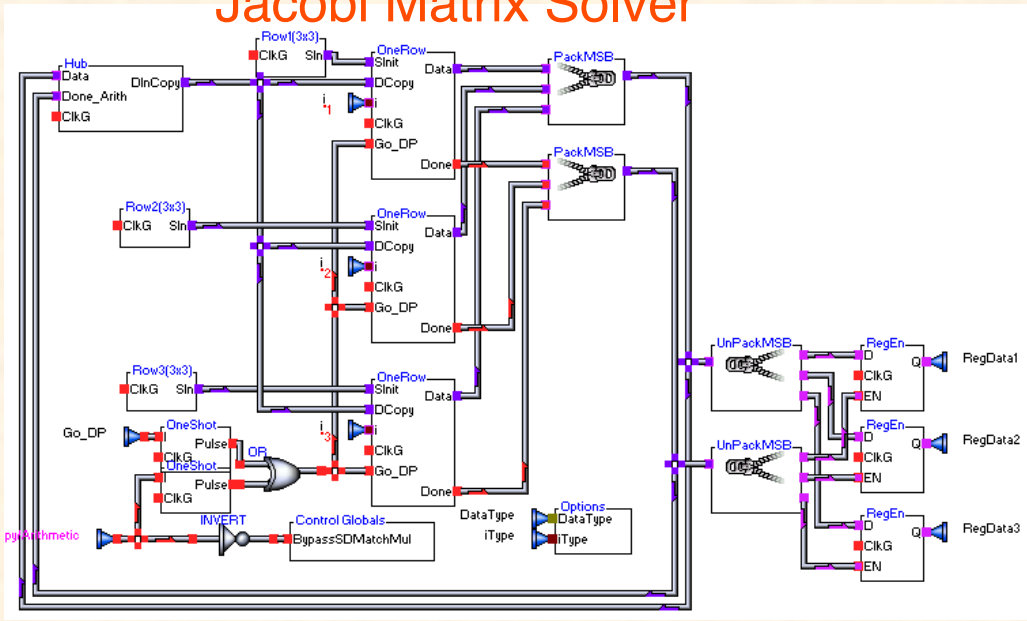
"write once, run anywhere"

Future Technologies Group ornl

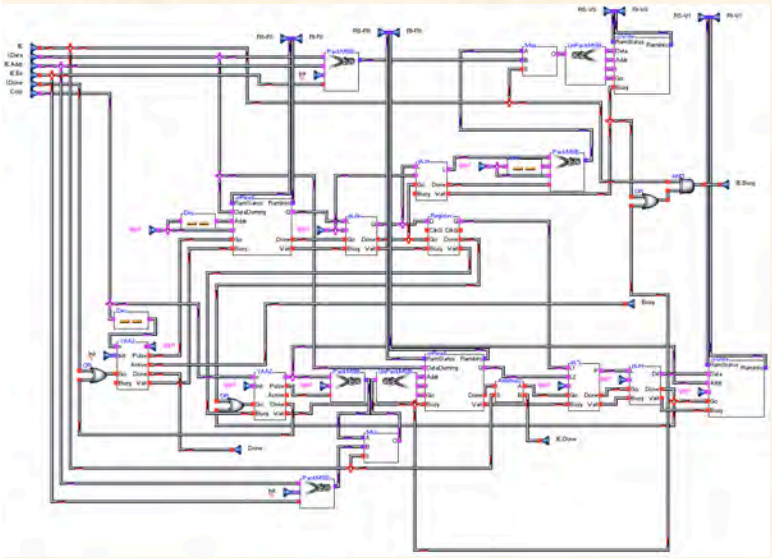
OAK RIDGE National Laboratory

Applications: VIVA Code

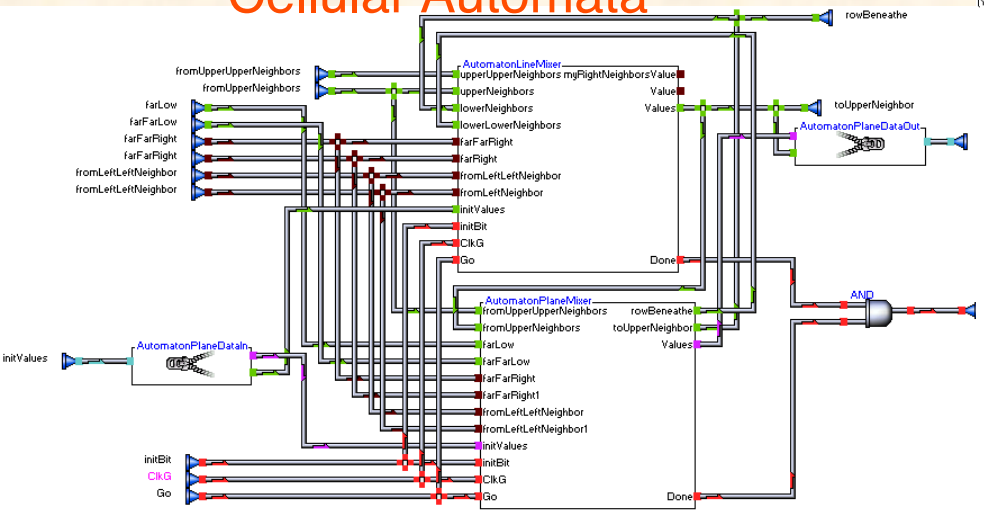
Jacobi Matrix Solver



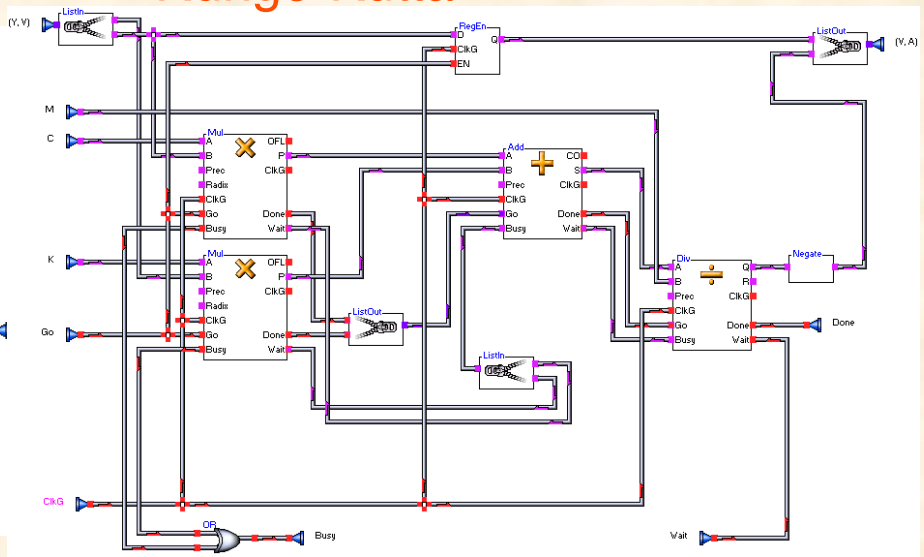
Gauss Matrix Solver



Cellular Automata



Runge-Kutta

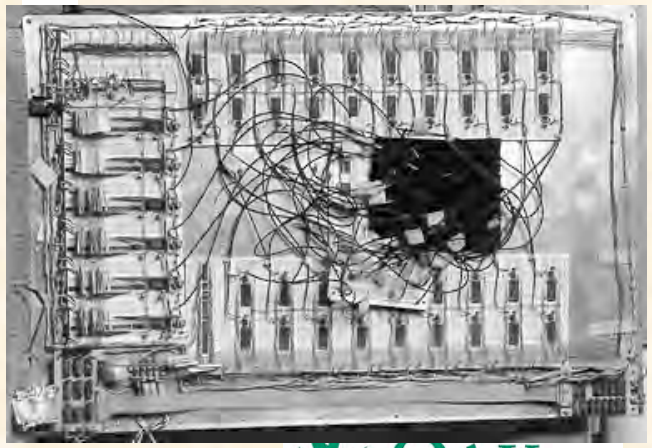


Future Technologies Group ornl

OAK RIDGE National Laboratory

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

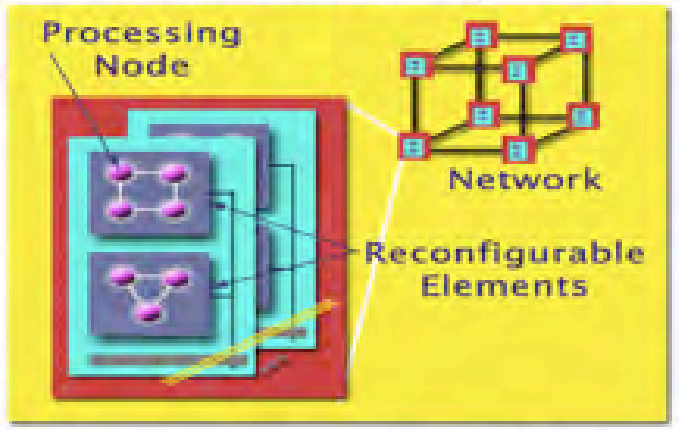


Future Technologies Group ornl

OAK RIDGE National Laboratory

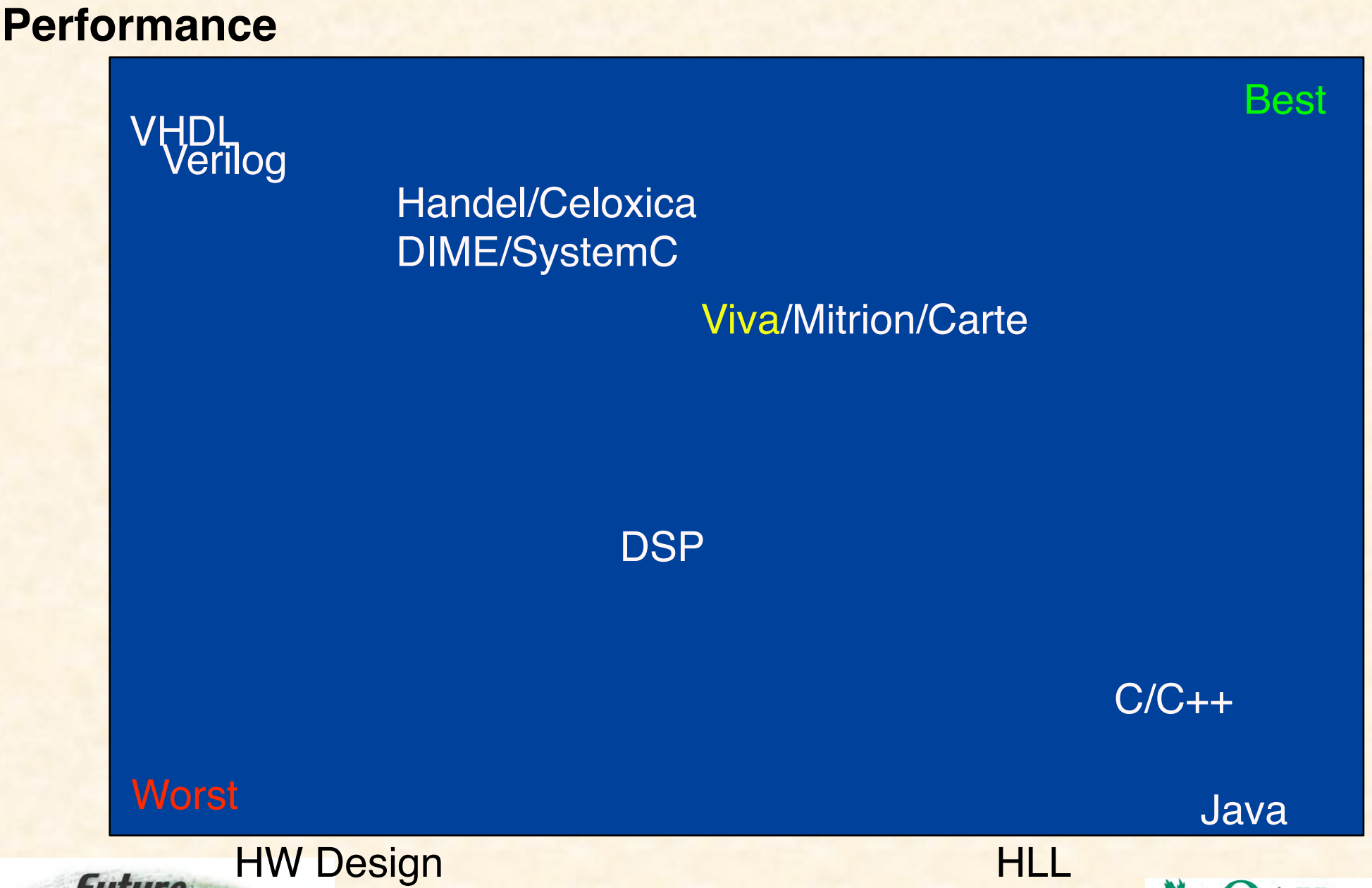
NASA Reconfigurable Scalable Computer

- Small, stackable modules
  - Proc, IO, Net,Power
  - Tailor for mission



- Scales by:
  - # nodes/module
  - # modules/stack
  - # stacks in system

Code Performance vs. Coding Ease



Future Technologies Group ornl

OAK RIDGE National Laboratory

Mitrion XD1 FPGA Config file

1. Write FPGA algorithm in Mitrion toolkit - <file>.mitc
2. Generate Mitrion processor core:
   mitrion -platform xd1vp50 -configure <file>.mitc
3. Copy mitrion_template to project
4. Copy mitrion.vhd from 2 to project subdirectory
   /src/80-00XX_Mitrion/hdl/user_app/
5. Run synthesis+P&R in project subdirectory
   /src/80-00XX_Mitrion/par/xc2vp50: 'make top.twr'
6. Timing top.twr OK?
7. Create XD1 FPGA configuration file:
   fcu -c top.bin ufphdr (top.bin created in step 5)

A. Code host program

B. Compile host.c:

cc -o <exec file> <host file>.c -lufp -lpthread -lrapl -D_REENTRANT

NOTE: Do 'module load xilinx' at ssh fpga (FPGA node login)

Mitrion Success

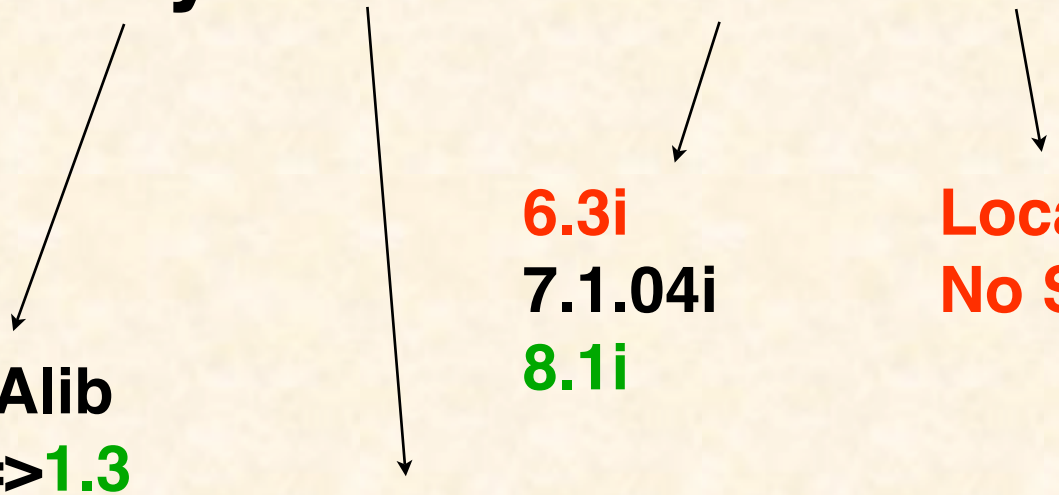
```
olaf@tiger-461-1:~/tiger> ./runhost
2381764 bytes loaded to FPGA
Resetting FPGA
Starting FPGA
Make sure mitrion processor is stopped
Have to clear bit0 of Reg 2 because Mitrion says so
Now we can start Mitrion
End by stopping mitrion processor
FPGA test completed
olaf@tiger-461-1:~/tiger> ll
```

?

Mitrion Matrix_add example:

ERROR:Place:543 - Due to placement constraints, the following 1 components cannot be placed. The relative offsets of the components are shown in brackets next to the component names.
FF user_app_inst_Mitrion_core_inst_lqmain_node_a_cutter_1566_Dsaved(10) (0, 0)

Cray-Mitrion-Xilinx-ORNL Issues



FPGAlib
1.2 =>1.3

Versions
0.98, 1.0.1, 1.1.1
1.1.2, 1.1.3
1.1.4, 1.1.5
mitrion_template missing

6.3i
7.1.04i
8.1i

Location
No Sysadmin

Future Technologies Group ornl

OAK RIDGE National Laboratory

Summary

- Cray XD1 with FPGAs
- Mitrion-C: *new* - ongoing evaluation
- SRC Carte & Viva - *success*
- Tools lag FPGA hardware =>
- No *silver bullet* limit Apps & HPC use *DARPA-hard* problem
- Promising FPGA future in HPC

Acknowledgements

- This research was sponsored by the Office of Mathematical, Information, and Computational Sciences, Office of Science, U.S. Department of Energy under Contract No. DE-AC05-00OR22725 with UT-Battelle, LLC. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

Cray
Dave Londo
Dave Strenski

Mitrion
Stefan, Anders
Pontus, Pex

ORNL
Chris Fuson

Ohio Supercomputer Center
Kevin Wohlever
Dennis Dalessandro

Future Technologies Group ornl

OAK RIDGE National Laboratory