**Document Info**

- **Title**: Reversible Logic for Supercomputing
- **Document Number**: 5231909
- **SAND Number**: 2005-2689 C
- **Review Type**: Electronic
- **Status**: Approved
- **Sandia Contact**: DEBENEDICTIS, ERIK P.
- **Submit Date**: 04/22/2005
- **Peer Reviewed?**: N

**Event (Conference/Journal/Book) Info**

- **Name**: Computing Frontiers 2005
- **City**: Ischia
- **State**: Country: Italy
- **Start Date**: 05/04/2005
- **End Date**: 05/06/2005

**Partnership Info**

- **Partnership Involved**: No
- **Partner Approval**: Agreement Number:

**Patent Info**

- **Scientific or Technical in Content**: Yes
- **Technical Advance**: No
- **TA Form Filed**: No

**Classification and Sensitivity Info**

- **Title**: Unclassified-Unlimited
- **Abstract**: Document: Unclassified-Unlimited
- **Additional Limited Release Info**: None.
- **DUSA**: None.

### Routing Details

<table>
<thead>
<tr>
<th>Role</th>
<th>Routed To</th>
<th>Approved By</th>
<th>Approval Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derivative Classifier Approver</td>
<td>SUMMERS, RANDALL M.</td>
<td>SUMMERS, RANDALL M.</td>
<td>04/22/2005</td>
</tr>
<tr>
<td>Role</td>
<td>Name</td>
<td>Second Name</td>
<td>Date</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>Classification Approver</td>
<td>WILLIAMS, RONALD L.</td>
<td>WILLIAMS, RONALD L.</td>
<td>04/25/2005</td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager Approver</td>
<td>PUNDIT, NEIL D.</td>
<td>PUNDIT, NEIL D.</td>
<td>04/29/2005</td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrator Approver</td>
<td>LUCERO, ARLENE M.</td>
<td>KRAMER, SAMUEL</td>
<td>06/05/2007</td>
</tr>
</tbody>
</table>

**Created by WebCo**  Problems? Contact CCHD: by email or at 845-CCHD (2243).

For Review and Approval process questions please contact the **Application Process Owner**
Reversible Logic for Supercomputing

How to save the Earth with Reversible Computing

Erik P. DeBenedictis
Sandia National Laboratories

May 5, 2005
Applications and $100M Supercomputers

**System Performance**

- 1 Zettaflops
- 100 Exaflops
- 10 Exaflops
- 1 Exaflops
- 100 Petaflops
- 10 Petaflops
- 1 Petaflops
- 100 Teraflops

**Applications**

- Plasma Fusion Simulation [Jardin 03]
- Full Global Climate [Malone 03]
- Compute as fast as the engineer can think [NASA 99]
- MEMS Optimize

**Technology**

- Nanotech + Reversible Logic [P]
  - (green) best-case logic
  - (red) worst-case logic
- Architecture: IBM Cyclops, FPGA, PIM
- Red Storm/Cluster

No schedule provided by source

---


[SCALeS 03] Workshop on the Science Case for Large-scale Simulation, June 24-25, proceedings on Internet at http://www.pnl.gov/scales/.

Objectives and Challenges

• Could reversible computing have a role in solving important problems?
  – Maybe, because power is a limiting factor for computers and reversible logic cuts power
• However, a complete computer system is more than “low power”
  – Processing, memory, communication in right balance for application
  – Speed must match user’s impatience
  – Must use a real device, not just an abstract reversible device
Outline

• An Exemplary Zettaflops Problem
• The Limits of Current Technology
• Arbitrary Architectures for the Current Problem
  – Searching the Architecture Space
  – Bending the Rules to Find Something
  – Exemplary Solution
• Conclusions
“Simulations of the response to natural forcings alone … do not explain the warming in the second half of the century”

“.model estimates that take into account both greenhouse gases and sulphate aerosols are consistent with observations over this period” - IPCC 2001
# FLOPS Increases for Global Climate

<table>
<thead>
<tr>
<th>Issue</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Zettaflops</td>
<td>Ensembles, scenarios</td>
</tr>
<tr>
<td></td>
<td>10×</td>
</tr>
<tr>
<td></td>
<td>Embarrassingly</td>
</tr>
<tr>
<td></td>
<td>Parallel</td>
</tr>
<tr>
<td>100 Exaflops</td>
<td>Run length</td>
</tr>
<tr>
<td></td>
<td>100×</td>
</tr>
<tr>
<td></td>
<td>Longer Running</td>
</tr>
<tr>
<td></td>
<td>Time</td>
</tr>
<tr>
<td>1 Exaflops</td>
<td>New parameterizations</td>
</tr>
<tr>
<td></td>
<td>100×</td>
</tr>
<tr>
<td></td>
<td>More Complex</td>
</tr>
<tr>
<td></td>
<td>Physics</td>
</tr>
<tr>
<td>10 Petaflops</td>
<td>Model Completeness</td>
</tr>
<tr>
<td></td>
<td>100×</td>
</tr>
<tr>
<td></td>
<td>More Complex</td>
</tr>
<tr>
<td></td>
<td>Physics</td>
</tr>
<tr>
<td>100 Teraflops</td>
<td>Spatial Resolution</td>
</tr>
<tr>
<td></td>
<td>$10^4 \times (10^3 \times 10^5 \times)$</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
</tr>
<tr>
<td>10 Gigaflops</td>
<td>Clusters Now In Use</td>
</tr>
<tr>
<td></td>
<td>(100 nodes, 5% efficient)</td>
</tr>
</tbody>
</table>

Outline

• An Exemplary Zettaflops Problem

• The Limits of Current Technology

• Arbitrary Architectures for the Current Problem
  – Searching the Architecture Space
  – Bending the Rules to Find Something
  – Exemplary Solution

• Conclusions
## Scientific Supercomputer Limits

<table>
<thead>
<tr>
<th>Expert Opinion</th>
<th>Best-Case Logic</th>
<th>Microprocessor Architecture</th>
<th>Physical Factor</th>
<th>Source of Authority</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 Exaflops</td>
<td>800 Petaflops</td>
<td>Reliability limit</td>
<td>Esteemed physicists</td>
</tr>
<tr>
<td></td>
<td>25 Exaflops</td>
<td>200 Petaflops</td>
<td>Derate 20,000 convert logic ops to floating point</td>
<td>Floating point engineering</td>
</tr>
<tr>
<td></td>
<td>4 Exaflops</td>
<td>32 Petaflops</td>
<td>Derate for manufacturing margin (4×)</td>
<td>Estimate</td>
</tr>
<tr>
<td></td>
<td>1 Exaflops</td>
<td>8 Petaflops</td>
<td>Uncertainty (6×)</td>
<td>Gap in chart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80 Teraflops</td>
<td>Improved devices (4×)</td>
<td>Estimate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40 Teraflops</td>
<td>Projected ITRS improvement to 22 nm (100×)</td>
<td>ITRS committee of experts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower supply voltage (2×)</td>
<td>ITRS committee of experts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Red Storm</td>
<td>contract</td>
</tr>
</tbody>
</table>

Assumption: Supercomputer is size & cost of Red Storm: US$100M budget; consumes 2 MW wall power; 750 KW to active components.

Derate 20,000 convert logic ops to floating point (64 bit precision).

Reliability limit 750KW/(80kB T)² (T=60°C junction temperature).
Outline

• An Exemplary Zettaflops Problem
• The Limits of Current Technology
• Arbitrary Architectures for the Current Problem
  – Searching the Architecture Space
  – Bending the Rules to Find Something
  – Exemplary Solution
• Conclusions
Supercomputer Expert System

Application/Algorithm run time model as in applications modeling

Logic & Memory Technology design rules and performance parameters for various technologies (CMOS, Quantum Dots, C Nano-tubes …)

Interconnect Speed, power, pin count, etc.

Physical Cooling, packaging, etc.

Expert System & Optimizer (looks for best 3D mesh of generalized MPI connected nodes, μP and other)

Time Trend Lithography as a function of years into the future

Results
1. Block diagram picture of optimal system (model)
2. Report of FLOPS count as a function of years into the future
Sample Analytical Runtime Model

- Simple case: finite difference equation
- Each node holds $n \times n \times n$ grid points

- Volume-area rule
  - Computing $\propto n^3$
  - Communications $\propto n^2$

$$T_{step} = 6 n^2 C_{bytes} T_{byte} + n^3 F_{grind/floprate}$$
Expert System for Future Supercomputers

- Applications Modeling
  - Runtime
    \[ T_{\text{run}} = f_1(n, \text{design}) \]
- Technology Roadmap
  - Gate speed = \( f_2(\text{year}) \),
  - chip density = \( f_3(\text{year}) \),
  - cost = \( $(n, \text{design}) \)
- Scaling Objective Function
  - I have \( C_1 \) & can wait \( T_{\text{run}} = C_2 \) seconds. What is the biggest \( n \) I can solve in year \( Y \)?

- Use “Expert System” To Calculate:
  \[ \max n: C_1 < n, T_{\text{run}} < C_2 \]
  All designs

- Report:
  \[ \text{Floating operations} \]
  \[ T_{\text{run}}(n, \text{design}) \]
  and illustrate “design”
Outline

• An Exemplary Zettaflops Problem
• The Limits of Current Technology
• Arbitrary Architectures for the Current Problem
  – Searching the Architecture Space
  – Bending the Rules to Find Something
  – Exemplary Solution
• Conclusions
The Big Issue

• Initially, didn’t meet constraints

- More Parallelism
  - Scaled Climate Model
    - 2D → 3D mesh, one cell per processor
  - Parallelize cloud-resolving model and ensembles

- More Device Speed
  - One Barely Plausible Solution
    - Consider special purpose logic with fast logic and low-power memory
  - Consider only highest performance published nanotech device QDCA

- Initial reversible nanotech
<table>
<thead>
<tr>
<th>Technology</th>
<th>Speed (min-max)</th>
<th>Dimension (min-max)</th>
<th>Energy per gate-op</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS</td>
<td>30 ps-1 μs</td>
<td>8 nm-5 μm</td>
<td>4 aJ</td>
<td></td>
</tr>
<tr>
<td>RSFQ</td>
<td>1 ps-50 ps</td>
<td>300 nm-1 μm</td>
<td>2 aJ</td>
<td>Larger</td>
</tr>
<tr>
<td>Molecular</td>
<td>10 ns-1 ms</td>
<td>1 nm-5 nm</td>
<td>10 zJ</td>
<td>Slower</td>
</tr>
<tr>
<td>Plastic</td>
<td>100 μs-1 ms</td>
<td>100 μm-1 mm</td>
<td>4 aJ</td>
<td>Larger+Slower</td>
</tr>
<tr>
<td>Optical</td>
<td>100 as-1 ps</td>
<td>200 nm-2 μm</td>
<td>1 pJ</td>
<td>Larger+Hotter</td>
</tr>
<tr>
<td>NEMS</td>
<td>100 ns-1 ms</td>
<td>10-100 nm</td>
<td>1 zJ</td>
<td>Slower+Larger</td>
</tr>
<tr>
<td>Biological</td>
<td>100 fs-100 μs</td>
<td>6-50 μm</td>
<td>.3 yJ</td>
<td>Slower+Larger</td>
</tr>
<tr>
<td>Quantum</td>
<td>100 as-1 fs</td>
<td>10-100 nm</td>
<td>1 zJ</td>
<td>Larger</td>
</tr>
<tr>
<td>QDCA</td>
<td>100 fs-10ps</td>
<td>1-10 nm</td>
<td>1 yJ</td>
<td>Smaller, faster, cooler</td>
</tr>
</tbody>
</table>

Data from ITRS ERD Section, data from Notre Dame
Outline

• An Exemplary Zettaflops Problem
• The Limits of Current Technology
• Arbitrary Architectures for the Current Problem
  – Searching the Architecture Space
  – Bending the Rules to Find Something
  – Exemplary Solution
• Conclusions
An Exemplary Device: Quantum Dots

- Pairs of molecules create a memory cell or a logic gate

Ref. “Clocked Molecular Quantum-Dot Cellular Automata,” Craig S. Lent and Beth Isaksen
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 50, NO. 9, SEPTEMBER 2003
1 Zettaflops Scientific Supercomputer

• How could we increase “Red Storm” from 40 Teraflops to 1 Zettaflops?
• Answer
  – >2.5×10^7 power reduction per operation
  – Faster devices × more parallelism >2.5×10^7
  – Smaller devices to fit existing packaging

A number of post-transistor devices have been proposed.

The shape of the performance curves have been validated by a consensus of reputable physicists.

However, validity of any data point can be questioned.

Cross-checking appropriate; see →


QCA Microprocessor Status

• M. Niemier Ph. D. Thesis, University of Notre Dame
• 12 Bit μP
• CAD design tool principles
  – 10× circuit density of CMOS at same λ
• Applies to various devices
  – Metal dot 4.2 nm²
  – Molecular 1.1 nm²

Figure 4.6. A 2-bit QCA Simple 12 ALU with Registers
Reversible Microprocessor Status

• Status
  – Subject of Ph. D. thesis
  – Chip laid out (no floating point)
  – RISC instruction set
  – C-like language
  – Compiler
  – Demonstrated on a PDE
  – However: really weird and not general to program with +=, -=, etc. rather than =
CPU Design

• Leading Thoughts
  – Implement CPU logic using reversible logic
    • High efficiency for the component doing the most logic
  – Implement state and memory using conventional logic
    • Low efficiency, but not many operations
  – Permits programming much like today
Atmosphere Simulation at a Zettaflops

Supercomputer is 211K chips, each with 70.7K nodes of 5.77K cells of 240 bytes; solves 86T=44.1Kx44.1Kx44.1K cell problem. System dissipates 332KW from the faces of a cube 1.53m on a side, for a power density of 47.3KW/m². Power: 332KW active components; 1.33MW refrigeration; 3.32MW wall power; 6.65MW from power company. System has been inflated by 2.57 over minimum size to provide enough surface area to avoid overheating. Chips are at 99.22% full, comprised of 7.07G logic, 101M memory decoder, and 6.44T memory transistors. Gate cell edge is 34.4nm (logic), 34.4nm (decoder); memory cell edge is 4.5nm (memory). Compute power is 768 EFLOPS, completing an iteration in 224µs and a run in 9.88s.
Performance Curve

FLOPS rate on Atmosphere Simulation →

Year →

Custom QDCA Rev. Logic

Rev. Logic Microprocessor

Custom

Cluster
Outline

• An Exemplary Zettaflops Problem
• The Limits of Current Technology
• Arbitrary Architectures for the Current Problem
  – Searching the Architecture Space
  – Bending the Rules to Find Something
  – Exemplary Solution
• Conclusions
Conclusions

- There are important applications that are believed to exceed the limits of irreversible logic
  - At US$100M budget
  - E.g. solution to global warming
- Reversible logic & nanotech point in the right direction
  - Low power

- Device Requirements
  - Push speed of light limit
  - Substantially sub-$k_B T$
  - Molecular scales
- Software and Algorithms
  - Must be much more parallel than today
- With all this, just barely works
- Conclusions appear to apply generally
Backup
<table>
<thead>
<tr>
<th>Expert Opinion</th>
<th>Best-Case Logic</th>
<th>Microprocessor Architecture</th>
<th>Physical Factor</th>
<th>Source of Authority</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 Exaflops</td>
<td>125:1</td>
<td>800 Petaflops</td>
<td>Reliability limit</td>
<td>Esteemed physicists</td>
</tr>
<tr>
<td>25 Exaflops</td>
<td>Derate 20,000 convert logic ops to floating point</td>
<td>200 Petaflops</td>
<td>Est. physicists (T=60°C junction temperature)</td>
<td></td>
</tr>
<tr>
<td>4 Exaflops</td>
<td>Derate for manufacturing margin (4×)</td>
<td>32 Petaflops</td>
<td>Est. physicists</td>
<td></td>
</tr>
<tr>
<td>1 Exaflops</td>
<td>Uncertainty (6×)</td>
<td>8 Petaflops</td>
<td>Gap in chart</td>
<td></td>
</tr>
<tr>
<td>1 Exaflops</td>
<td>Improved devices (4×)</td>
<td>8 Petaflops</td>
<td>Est. physicists</td>
<td></td>
</tr>
<tr>
<td>8 Exaflops</td>
<td>Improved devices (4×)</td>
<td>80 Teraflops</td>
<td>Est. physicists</td>
<td></td>
</tr>
<tr>
<td>80 Teraflops</td>
<td>Projected ITRS improvement to 22 nm (100×)</td>
<td>40 Teraflops</td>
<td>ITRS committee of experts</td>
<td></td>
</tr>
<tr>
<td>40 Teraflops</td>
<td>Lower supply voltage (2×)</td>
<td>40 Teraflops</td>
<td>ITRS committee of experts</td>
<td></td>
</tr>
</tbody>
</table>

Assumption: Supercomputer is size & cost of Red Storm: US$100M budget; consumes 2 MW wall power; 750 KW to active components.
Metaphor: FM Radio on Trip to in USA

• You drive to a distant listening to FM radio

• Music clear for a while, but noise creeps in and then overtakes music

• Analogy: You live out the next dozen years buying PCs every couple years

• PCs keep getting faster
  – clock rate increases
  – fan gets bigger
  – won’t go on forever

• Why…see next slide

Driving away from FM transmitter $\rightarrow$ less signal
Noise from electrons $\rightarrow$ no change

Increasing numbers of gates $\rightarrow$ less signal power
Noise from electrons $\rightarrow$ no change