Review Approval

- Prepare Request
- Search Requests
- Generate Reports
- Approvals
- Help
  - Wizard

Search Requests
- New Search
- Refine Search
- Search Results

- Clone Request
- Edit Request
- Cancel Request

Search Detail

Submittal Details

Document Info

Title : Extreme Supercomputing

Document Number : 5225718 — SAND Number : 2004-4707 P

Review Type : Electronic — Status : Approved

Sandia Contact : DEBENEDICTIS,ERIK P. — Submittal Type : Viewgraph/Presentation

Requestor : DEBENEDICTIS,ERIK P. — Submit Date : 09/14/2004

Author(s)

DEBENEDICTIS,ERIK P.

Event (Conference/Journal/Book) Info

Name : Seminar at University of Notre Dame, CS & Eng. and Center for Nano Science and Technology

City : South Bend — State : IN — Country : USA

Start Date : 09/13/2004 — End Date : 09/13/2004

Partnership Info

Partnership Involved : No

Partner Approval : — Agreement Number :

Patent Info

Scientific or Technical in Content : Yes

Technical Advance : No — TA Form Filed : No

SD Number :

Classification and Sensitivity Info

Title : Unclassified-Unlimited — Abstract : — Document : Unclassified-Unlimited

Additional Limited Release Info : None.

DUSA : None.

Routing Details

Role	Routed To	Approved By	Approval Date
Derivative Classifier Approver	YARRINGTON,PAUL	YARRINGTON,PAUL	09/14/2004
Conditions:			

Classification Approver	WILLIAMS,RONALD L.	WILLIAMS,RONALD L.	**09/16/2004**
Conditions:			
Manager Approver	PUNDIT,NEIL D.	PUNDIT,NEIL D.	**09/17/2004**
Conditions:			
Administrator Approver	LUCERO,ARLENE M.	FARRELLY,JEREMIAH	**05/24/2007**

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**

SAND2004-4707P

Extreme Supercomputing

Erik P. DeBenedictis

Sandia National Laboratories

Presentation at University of Notre Dame

September 13, 2004

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

The Baseline

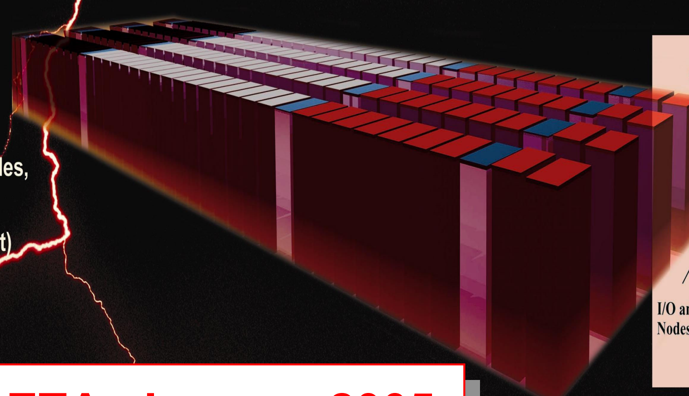
RED STORM

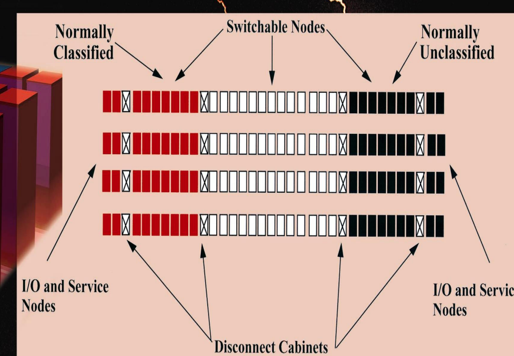
Design Parameters

- True MPP, Designed to be a single system
- Fully connected high performance 3-D mesh interconnect
- Topology - 27 X 16 X 24 compute nodes and 2 X 8X 16 service and I/O nodes
- 108 compute node cabinets and 10,368 compute node processors (AMD Sledgehammer @ 2.0 GHz)
- ~10 TB of DDR memory @ 333 MHz (1.0 GB per processor)
- Red/Black switching - ~1/4, ~1/2, ~1/4
- 8 Service and I/O cabinets on each end (256 processors for each color)
- 240 TB of disk storage (120 TB per color)
- Functional hardware partitioning - service and I/O nodes, compute nodes, and RAS nodes
- Functional system software partitioning - LINUX on service and I/O nodes, LWK (Catamount) on compute nodes, stripped down LINUX on RAS nodes
- Separate RAS and system management network (Ethernet)
- Router table based routing in the interconnect
- Less than 2 MW total power and cooling
- Less than 3,000 square feet of floor space

Performance

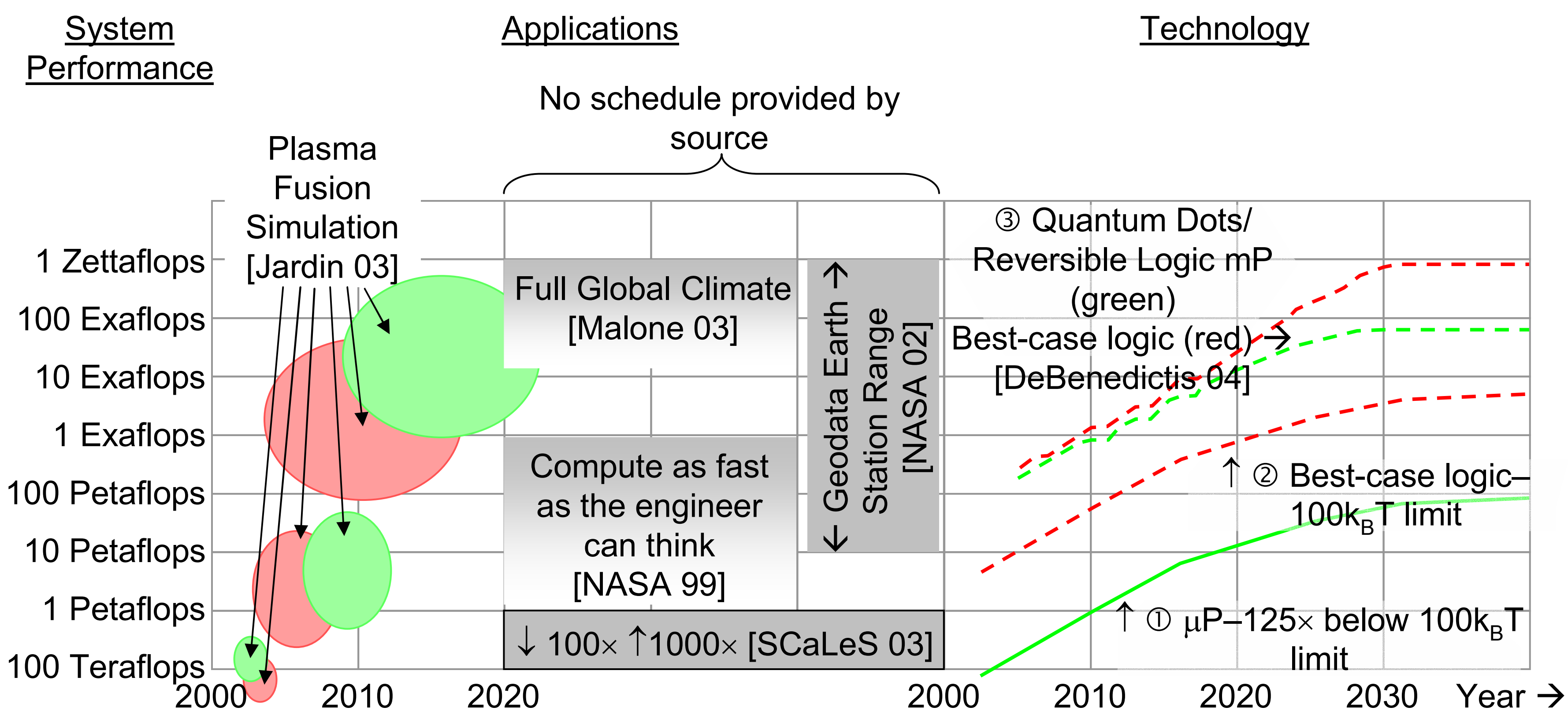
- Peak of ~ 40 TF
- Expected MP-Linpack performance >20 TF
- Aggregate system memory bandwidth ~55 TB/s
- Interconnect
  - Aggregate sustained interconnect bandwidth > 100 TB/s
  - MPI Latency - µs neighbor, 5 µs across machine
  - Bi-Section bandwidth ~2.3 TB/s
  - Link bandwidth ~3.0 GB/s in each direction
- I/O System
  - Sustained 50 GB/s disk I/O bandwidth for each color
  - Sustained 25 GB/s external network bandwidth for each color





ETA: January 2005

Applications and Computer Technology



[Jardin 03] S.C. Jardin, "Plasma Science Contribution to the SCaLeS Report," Princeton Plasma Physics Laboratory, PPPL-3879 UC-70, available on Internet.
[Malone 03] Robert C. Malone, John B. Drake, Philip W. Jones, Douglas A. Rotman, "High-End Computing in Climate Modeling," contribution to SCaLeS report.
[NASA 99] R. T. Biedron, P. Mehrotra, M. L. Nelson, F. S. Preston, J. J. Rehder, J. L. Rogers, D. H. Rudy, J. Sobieski, and O. O. Storaasli, "Compute as Fast as the Engineers Can Think!" NASA/TM-1999-209715, available on Internet.
[NASA 02] NASA Goddard Space Flight Center, "Advanced Weather Prediction Technologies: NASA's Contribution to the Operational Agencies," available on Internet.
[SCaLeS 03] Workshop on the Science Case for Large-scale Simulation, June 24-25, proceedings on Internet a http://www.pnl.gov/scales/.
[DeBenedictis 04], Erik P. DeBenedictis, "Matching Supercomputing to Progress in Science," July 2004. Presentation at Lawrence Berkeley National Laboratory, also published as Sandia National Laboratories SAND report SAND2004-3333P. Sandia technical reports are available by going to http://www.sandia.gov and accessing the technical library.

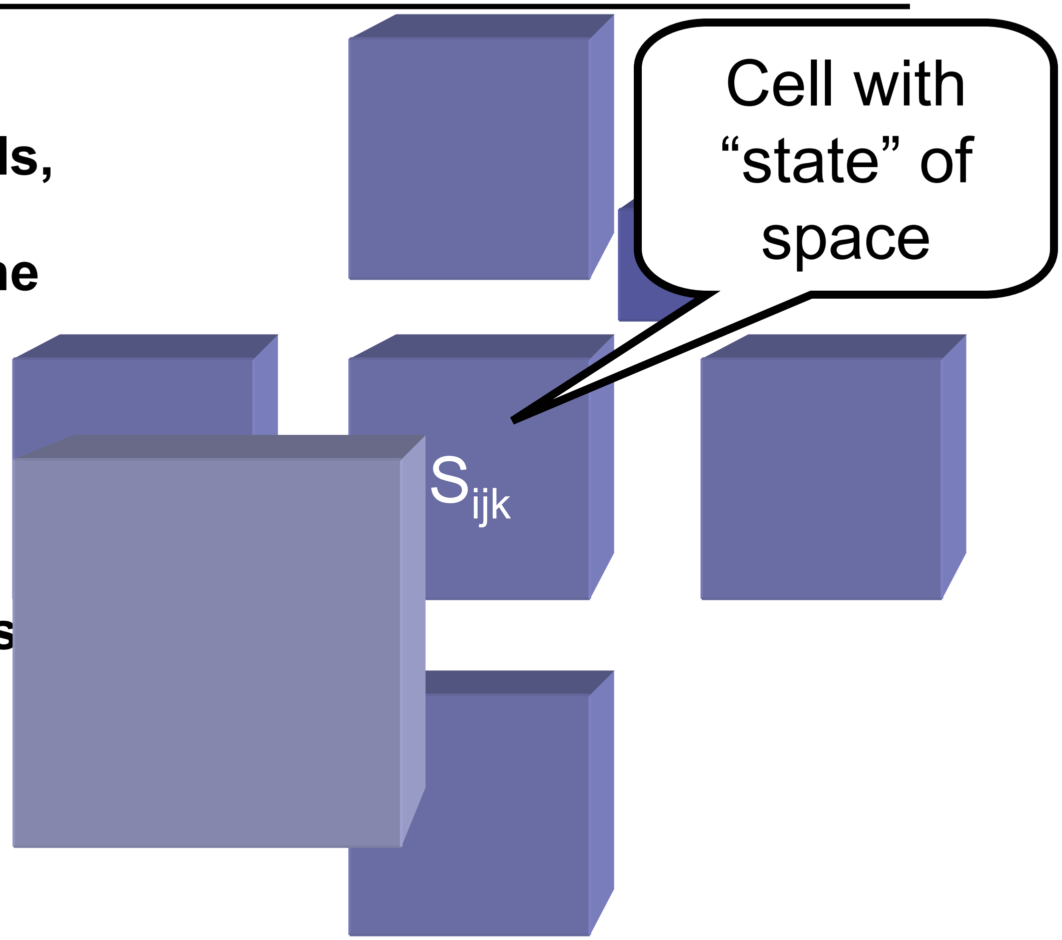
Sandia National Laboratories

Outline

- The Computing of Physics:
  The Need for Zettaflops
- Limits of Moore's Law
  Today's Technologies
- An Expert System/Optimizer for Supercomputing
- The Physics of Computing:
  Reaching to Zettaflops
- Roadmap and Future Directions

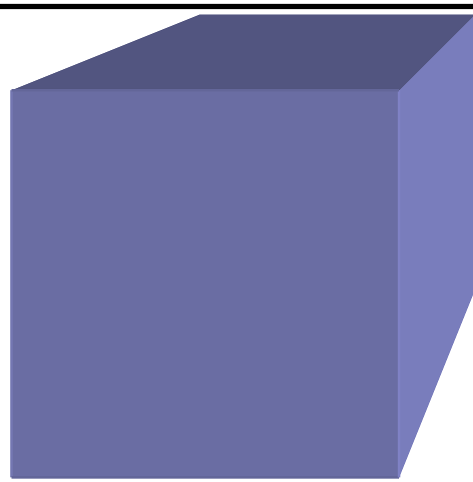
Simulation of Physics on a Computer

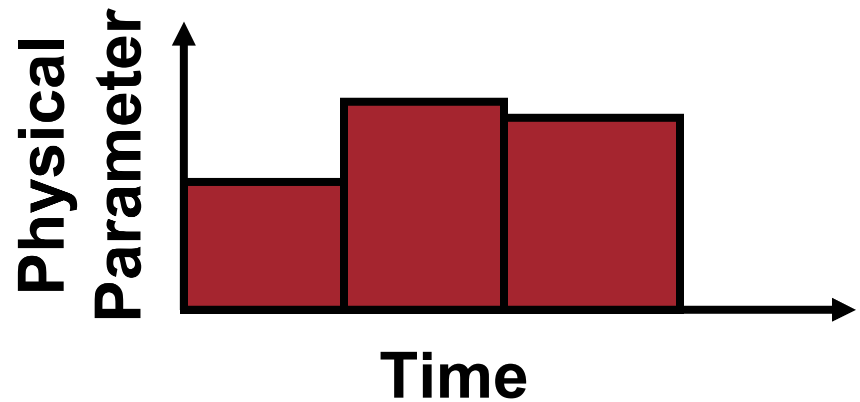
- **Space is divided into cells, each with computer variables representing the physical state of the volume represented by the cell**
- **The computer updates the state of a cell for successive time intervals ΔT based on some physical laws**
- **I. e. $S_{ijk}' = f(S_{ijk}$, states of nearby cells)**



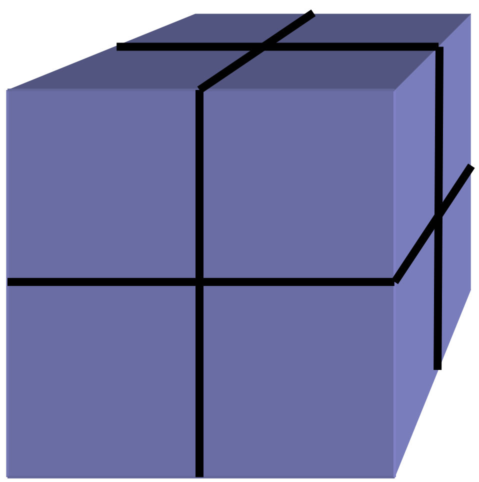
Fourth Power Scaling Rule

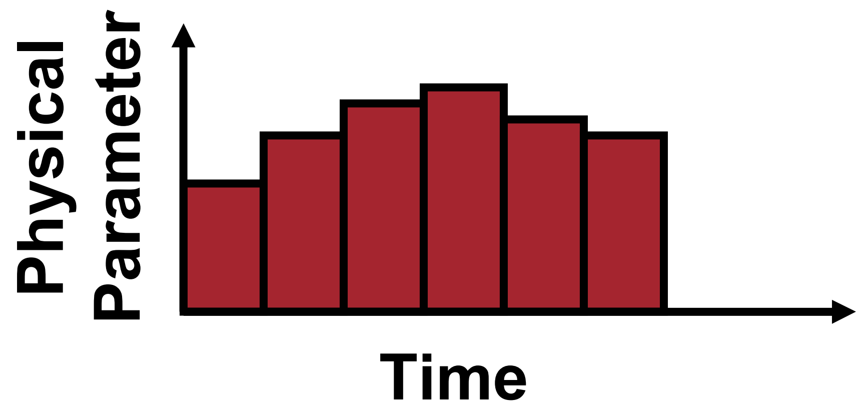
Reference





2× spatial resolution, 2× time steps → 2^4× FLOPS





Global Climate

- Objective
  - Collect data about Earth
  - Model climate into the future
  - Provide "decision support" and ability to "mitigate"
- Approaches
  - Climate models exist, but need they more resolution, better physics, and better initial conditions (observations of the Earth)
- Computer Resources Required
  - Increments over current workstation on next slide

FLOPS Increases for Global Climate

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

Issue	Scaling
Ensembles, scenarios 10×	Embarrassingly Parallel
Run length 100×	Longer Running Time
New parameterizations 100×	More Complex Physics
Model Completeness 100×	More Complex Physics
Spatial Resolution $10^4\times$ ($10^3\times$-$10^5\times$)	Resolution
Clusters Now In Use (100 nodes, 5% efficient)	

Ref. "High-End Computing in Climate Modeling," Robert C. Malone, LANL, John B. Drake, ORNL, Philip W. Jones, LANL, and Douglas A. Rotman, LLNL (2004)

Sandia National Laboratories

NASA Climate Earth Station

Based on these inputs, various portions of the Modeling and Data Assimilation System will require anywhere from 10^7 to 10^{13} GFLOPS of computational resources. In other words, the range of computational resources needed is 10^{16} to 10^{21} Floating Point Operations per Second. For the curious, the range can also be stated as 10 PetaFLOPS to 1 ZettaFLOPS.

4.1.2. Anticipated Computing Technology Capabilities

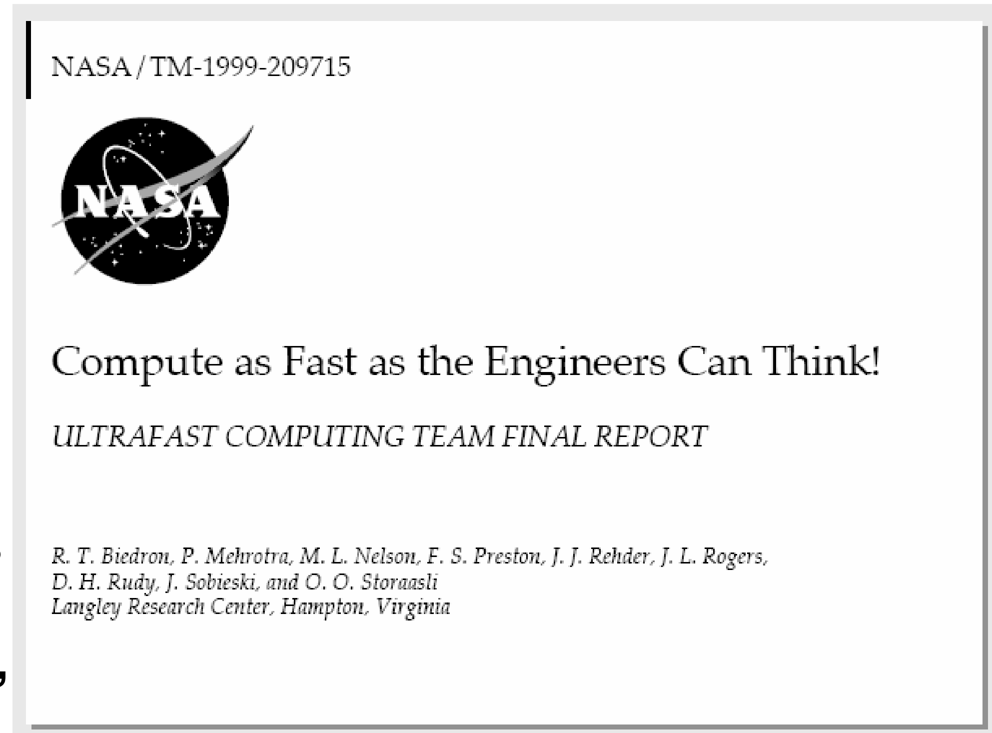
At first glance, the numbers discussed in the previous section appear so high as to be impossibly ludicrous. However, with the expected growth in computing capabilities, the lower end of this spectrum actually falls within the domain of possibility.

– **"Advanced Weather Prediction Technologies: NASA's Contribution to the Operational Agencies," Gap Analysis Appendix, May 31, 2002**

Sandia National Laboratories

NASA Work Station

- "…the ultimate goal of making the computing underlying the design process so capable that it no longer acts as a brake on the flow of the creative human thought…"
- Requirement 3 Exaflops
- Note: In the context of this report, this requirement is for one or a few engineers, not a supercomputer center!


NASA / TM-1999-209715

Compute as Fast as the Engineers Can Think!

ULTRAFAST COMPUTING TEAM FINAL REPORT

R. T. Biedron, P. Mehrotra, M. L. Nelson, F. S. Preston, J. J. Rehder, J. L. Rogers, D. H. Rudy, J. Sobieski, and O. O. Storaasli
Langley Research Center, Hampton, Virginia

Outline

- The Computing of Physics:
  The Need for Zettaflops
- Limits of Moore's Law
  Today's Technologies
- An Expert System/Optimizer for Supercomputing
- The Physics of Computing:
  Reaching to Zettaflops
- Roadmap and Future Directions

*** This is a Preview ***

Best-Case Logic	Microprocessor Architecture	Physical Factor	Source of Authority
2×10^{24} logic ops/s		Reliability limit 750KW/(80k_BT)	Esteemed physicists (T=60°C junction temperature)
		Derate 20,000 convert logic ops to floating point	Floating point engineering (64 bit precision)
Expert Opinion: 100 Exaflops	800 Petaflops	Derate for manufacturing margin (4×)	Estimate
← 125:1 →			
Estimate: 25 Exaflops	200 Petaflops	Uncertainty (6×)	Gap in chart
4 Exaflops	32 Petaflops	Improved devices (4×)	Estimate
1 Exaflops	8 Petaflops	Projected ITRS improvement to 22 nm (100×)	ITRS committee of experts
	80 Teraflops	Lower supply voltage (2×)	ITRS committee of experts
	40 Teraflops	Red Storm	contract

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

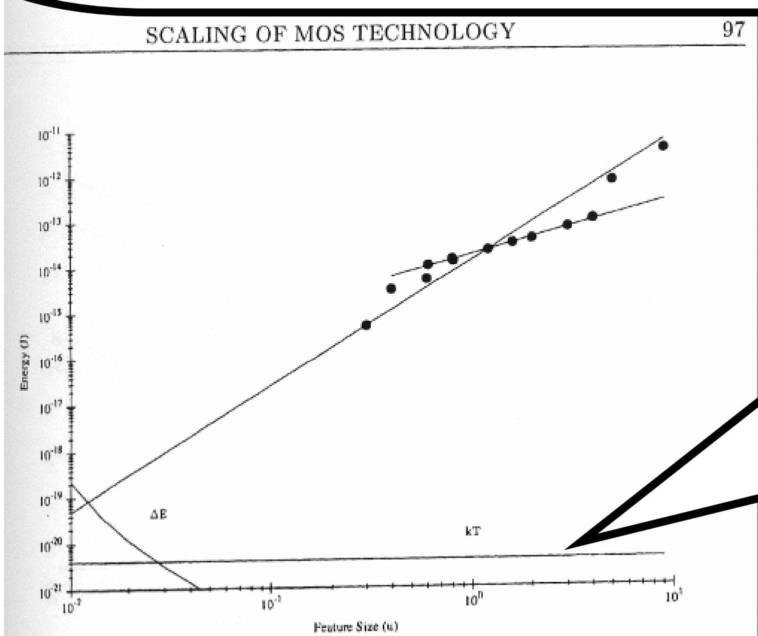
Sandia National Laboratories

Thermal Noise Limit

This logical irreversibility is associated with physical irreversibility and requires a minimal heat generation, per machine cycle, typically of the order of kT for each irreversible function.
– R. Landauer 1961


R. Landauer

Irreversibility and Heat Generation in the Computing Process



kT "helper line," drawn out of the reader's focus because it wasn't important at the time of writing
– Carver Mead, Scaling of MOS Technology, 1994

Sandia National Laboratories

Metaphor: FM Radio on Trip to Chicago

- You drive to Chicago 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

Details: Erik DeBenedictis, "Taking ASCI Supercomputing to the End Game," SAND2004-0959

SIA Semiconductor Roadmap

- Generalization of Moore's Law
  - Projects many parameters
  - Years through 2016
  - Includes justification
  - Panel of experts
    - known to be wrong
  - Size between Albuquerque white and yellow pages

International Technology Roadmap for Semiconductors (ITRS), see http://public.itrs.net

Sandia National Laboratories

Semiconductor Roadmap

YEAR OF PRODUCTION	2010	2013	2016
DRAM ½ PITCH (nm)	45	32	22
MPU / ASIC ½ PITCH (nm)	50	35	25
MPU PRINTED GATE LENGTH (nm)	25	18	13
MPU PHYSICAL GATE LENGTH (nm)	18	13	9
Physical gate length high-performance (HP) (nm) [1]	18	13	9
Equivalent physical oxide thickness for high-performance T_{ox} (EOT)(nm) [2]	0.5-0.8	0.4-0.6	0.4-0.5
Gate depletion and quantum effects electrical thickness adjustment factor (nm) [3]	0.5	0.5	0.5
T_{ox} electrical equivalent (nm) [4]	1.2	1.0	0.9
Nominal power supply voltage (V_{dd}) (V) [5]	0.6	0.5	0.4
Nominal high-performance NMOS sub threshold leakage current, $I_{sd,leak}$ (at 25 °C) (μA/μm) [6]	3	7	10
Nominal high-performance NMOS saturation drive current , I_{dd} (at V_{dd}, at 25 °C) (μA/μm) [7]	1200	1500	1500
Required percent current-drive "mobility/transconductance improvement" [8]	30%	70%	100%
Parasitic source/drain resistance (Rsd) (oh…	110	90	80
Parasitic source/drain resistance (Rsd) pe…	25%	30%	35%
Parasitic capacitance percent of ideal gat…	31%	36%	42%
High-performance NMOS device τ (C_{gate} * …	0.39	0.22	0.15
Relative device performance [13]	[illegible]	7.2	10.7
Energy per (W/L_{gate}=3) device switching transition (C_{gate}*(3*L_{gate})*V^2) (fJ/Device) [14]	0.015	0.007	0.002
Static power dissipation per (W/Lgate=3) device (Watts/Device) [15]	9.7E-08	1.4E-07	1.1E-07

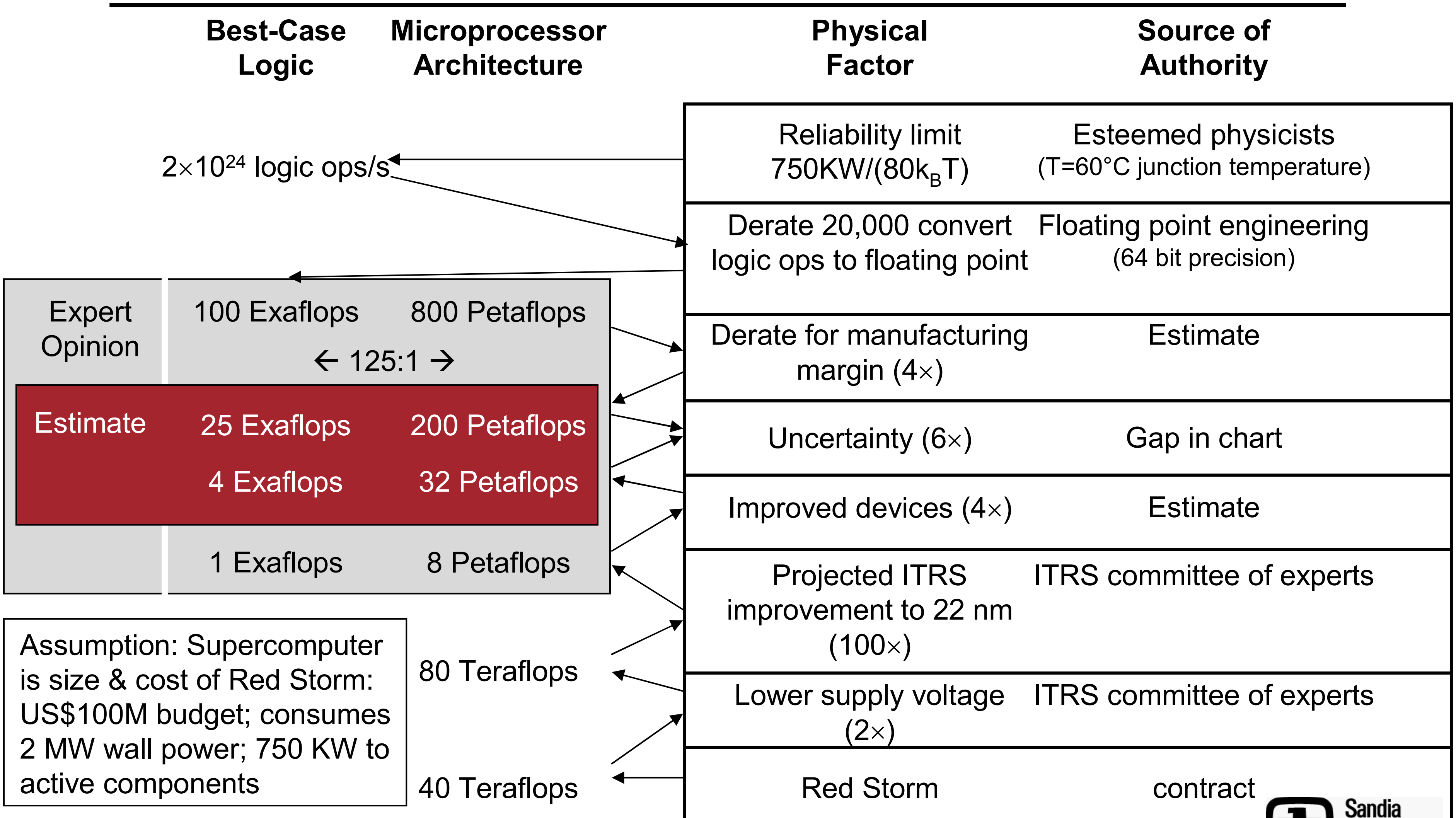
1,000 k_BT/transistor

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known

Scientific Supercomputer Limits



Sandia National Laboratories

Personal Observational Evidence

- Have radios become better able to receive distant stations over the last few decades with a rate of improvement similar to Moore's Law?

- You judge from your experience, but the answer should be that they have not.

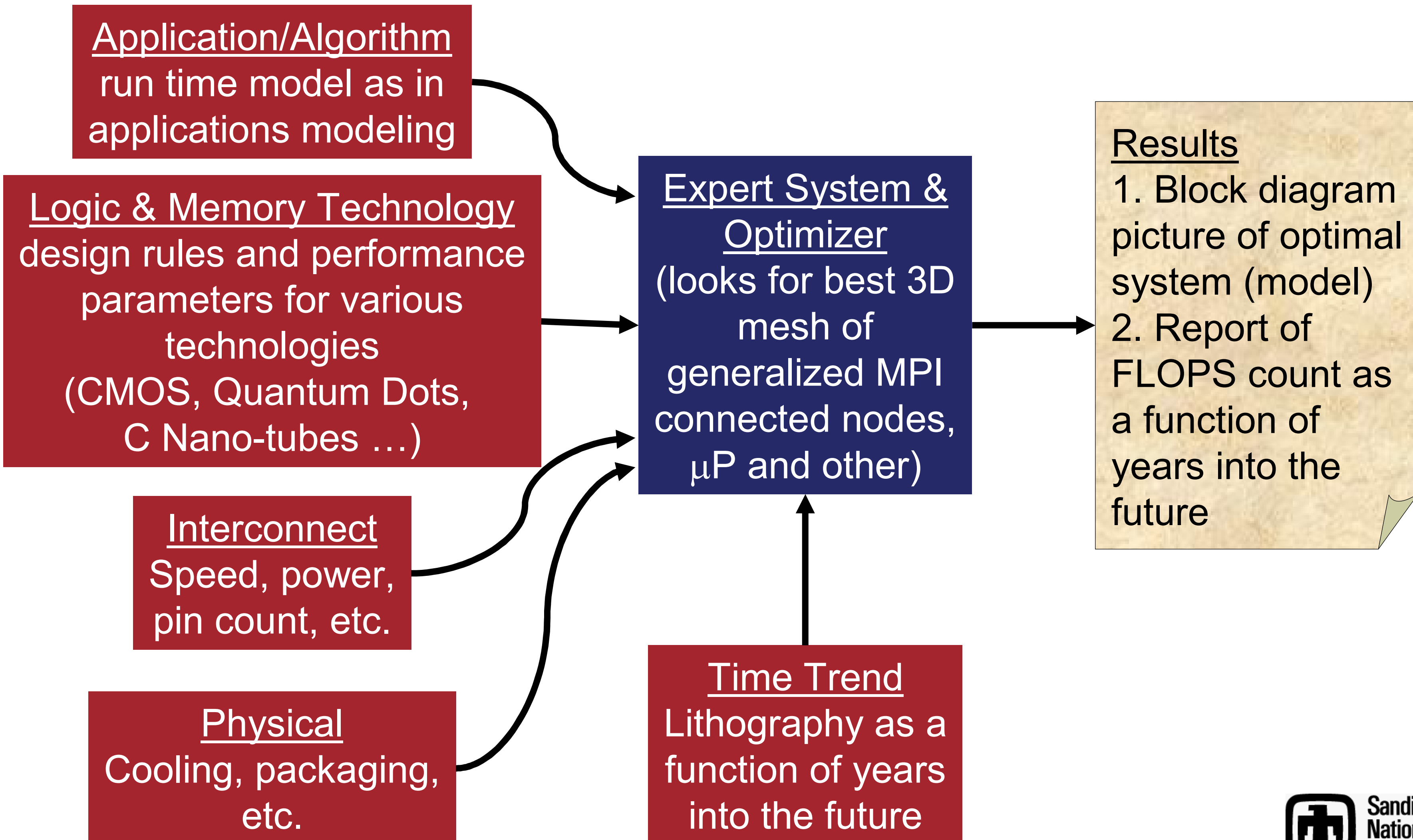
- Therefore, electrical noise does not scale with Moore's Law.

Outline

- The Computing of Physics:
  The Need for Zettaflops
- Limits of Moore's Law
  Today's Technologies
- An Expert System/Optimizer for Supercomputing
- The Physics of Computing:
  Reaching to Zettaflops
- Roadmap and Future Directions

Sandia National Laboratories

Supercomputer Expert System

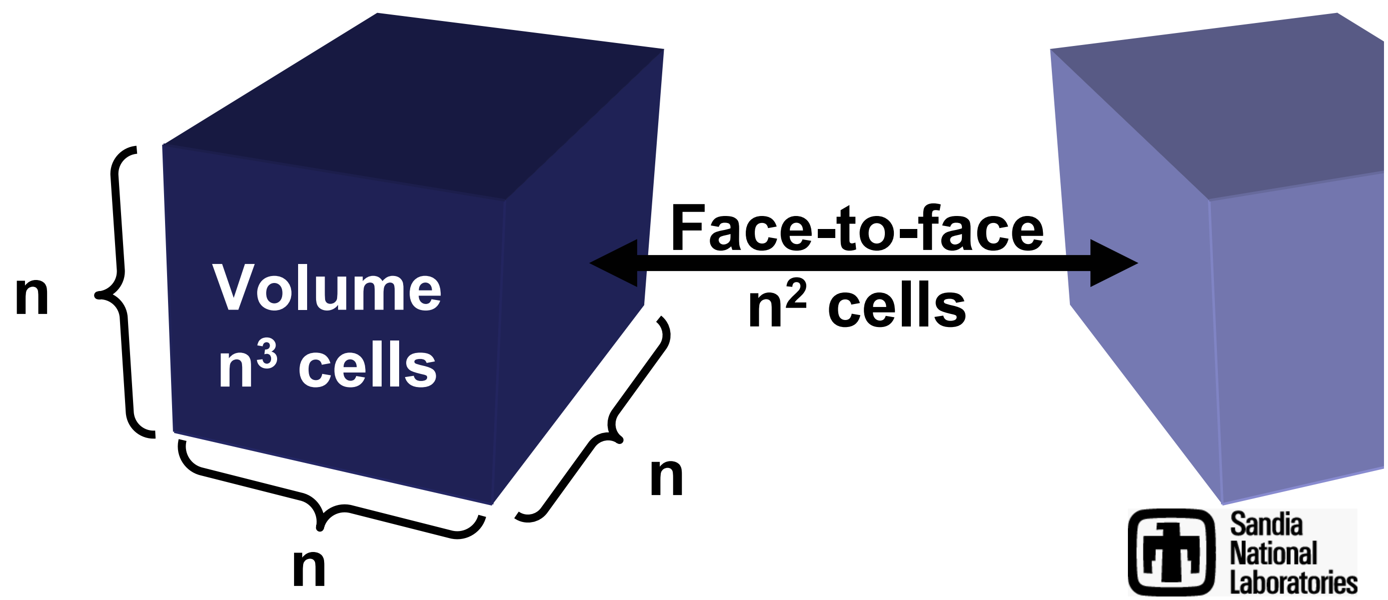


Sample Analytical Runtime Model

- Simple case: finite difference equation
- Each node holds n×n×n grid points

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

$$T_{step} = 6\, n^2\, C_{bytes}\, T_{byte} + n^3\, F_{grind}/\text{floprate}$$



Sandia National Laboratories

Expert System for Future Supercomputers

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

- Use "Expert System" To Calculate:

$$\max_{\text{All designs}} \quad n: \$<C_1,\ T_{run}<C_2$$

- Report:

$$\frac{\text{Floating operations}}{T_{run}(n, \text{design})}$$

and illustrate "design"

Sandia National Laboratories

Outline

- The Computing of Physics:
  The Need for Zettaflops
- Limits of Moore's Law
  Today's Technologies
- An Expert System/Optimizer for Supercomputing
- The Physics of Computing:
  Reaching to Zettaflops
- Roadmap and Future Directions

Sandia National Laboratories

Candidate Technologies for Zettaflops

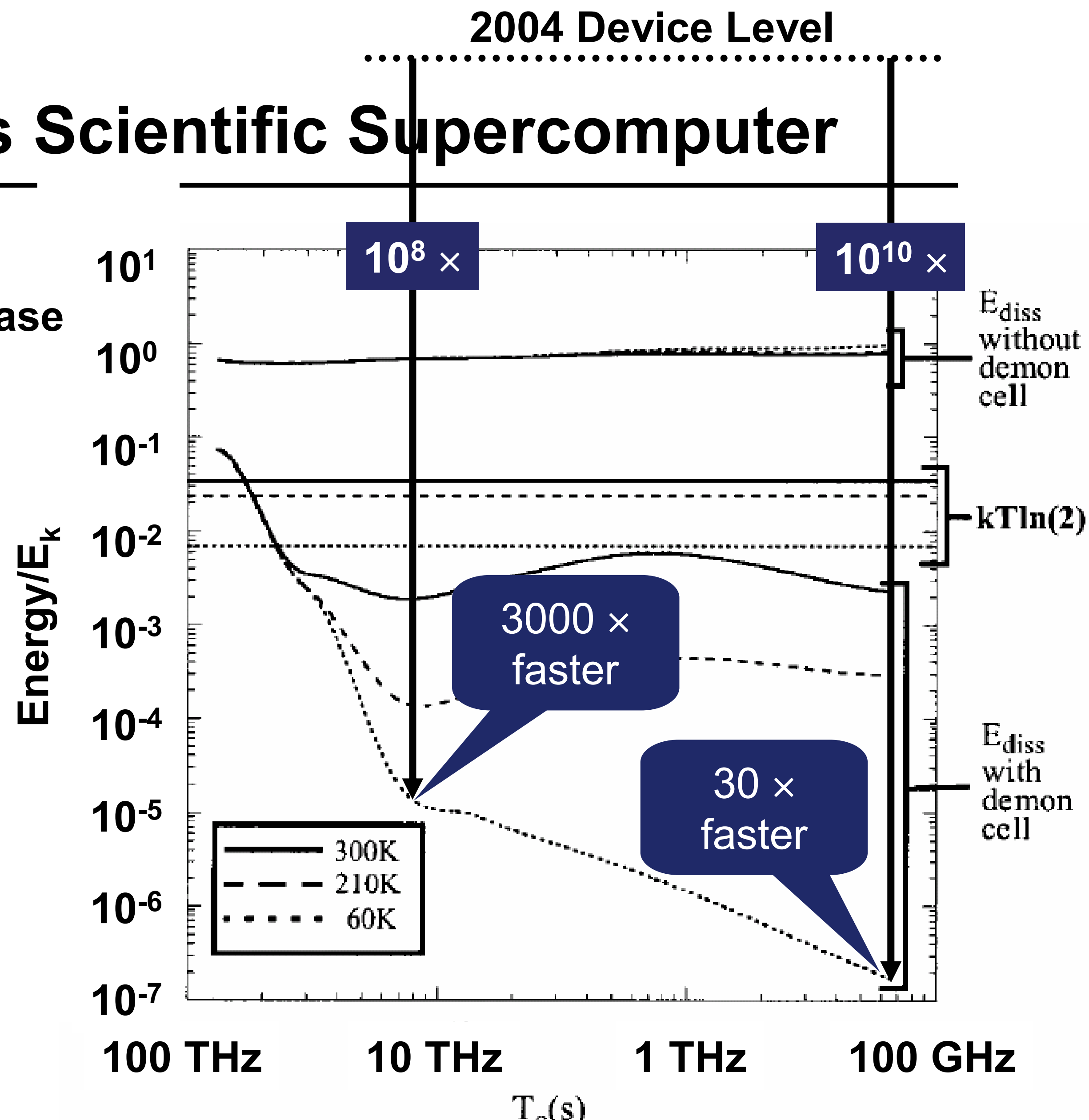
- CMOS per Moore's Law
  - Cluster/μP solution exceeds limits by 10,000×
    - Trillion US$ cost
    - 10 × Hoover Dam for power supply
  - Custom logic solution exceeds limits by 100×
    - US$10 billion cost
    - 100 MW power
  - ∴ worth our while to consider alternatives
- Limiting search for Alternatives to CMOS
  - Digital (not Analog)
  - Plausible to lots of Floating point
  - Controllable by something recognizable as "programming"
  - Mature enough for above issues to be addressed in published papers
  - Rules out coherent quantum, neural nets, DNA computing, optical interference, …

Alternatives to CMOS for Zettaflops

- New Devices
  - Superconducting: RSFQ (a. k. a. nSQUID, parametric quantrons)
  - Quantum Dots/QCA
  - Rod Logic
  - Helical Logic
  - Single Electron Transistors
  - Carbon Nanotube Y Junctions
  - …
- Logic and Architecture
  - "Reversible logic" will be unfamiliar to today's engineers but has been shown to be sufficient
  - Arithmetic elements and microprocessors have been demonstrated
  - Leading architecture:
    - Reversible ALU/CPU
    - Irreversible memory

1 Zettaflops Scientific Supercomputer

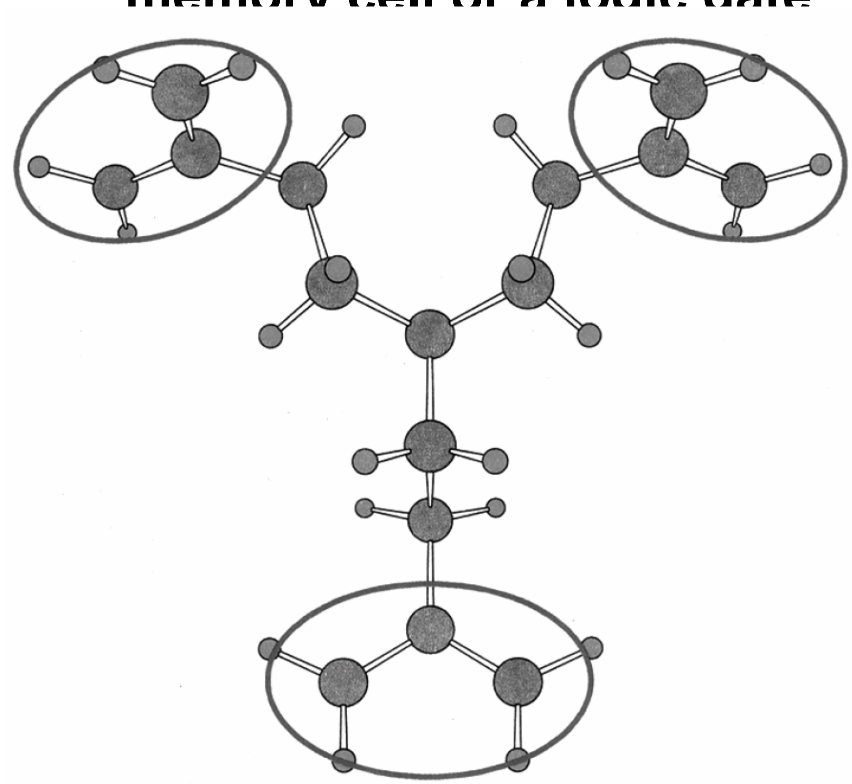
- How could we increase "Red Storm" from 40 Teraflops to 1 Zettaflops?
- Answer
  - >2.5×10⁷ power reduction per operation
  - Faster devices × more parallelism >2.5×10⁷
  - Smaller devices to fit existing packaging

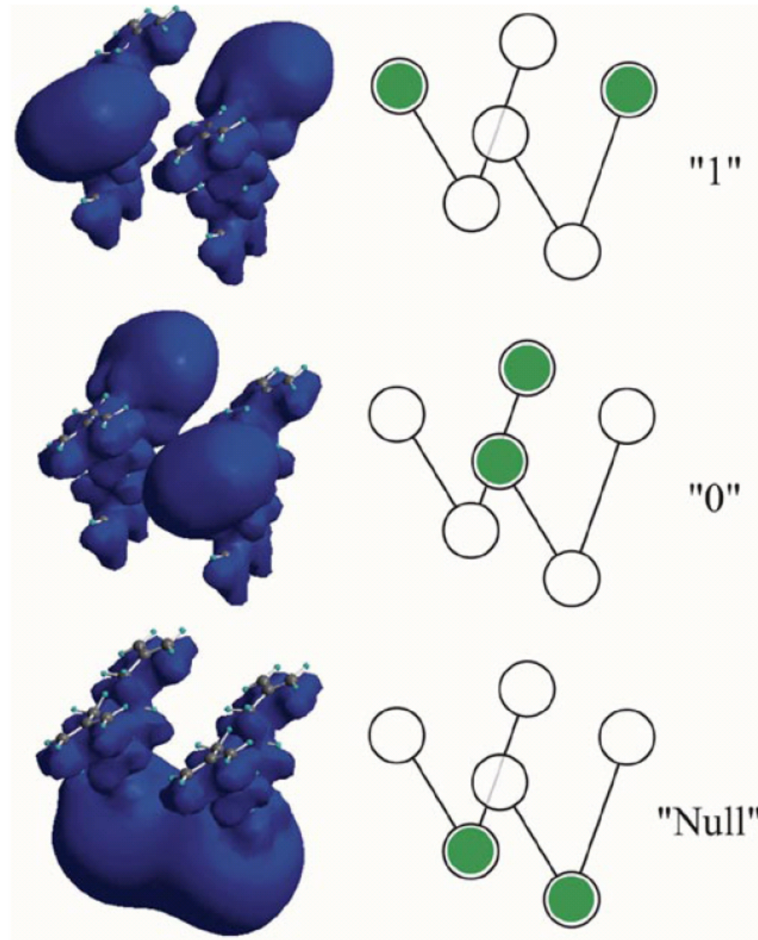


Ref. "Maxwell's demon and quantum-dot cellular automata," John Timler and Craig S. Lent, JOURNAL OF APPLIED PHYSICS 15 JULY 2003

An Exemplary Device: Quantum Dots

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

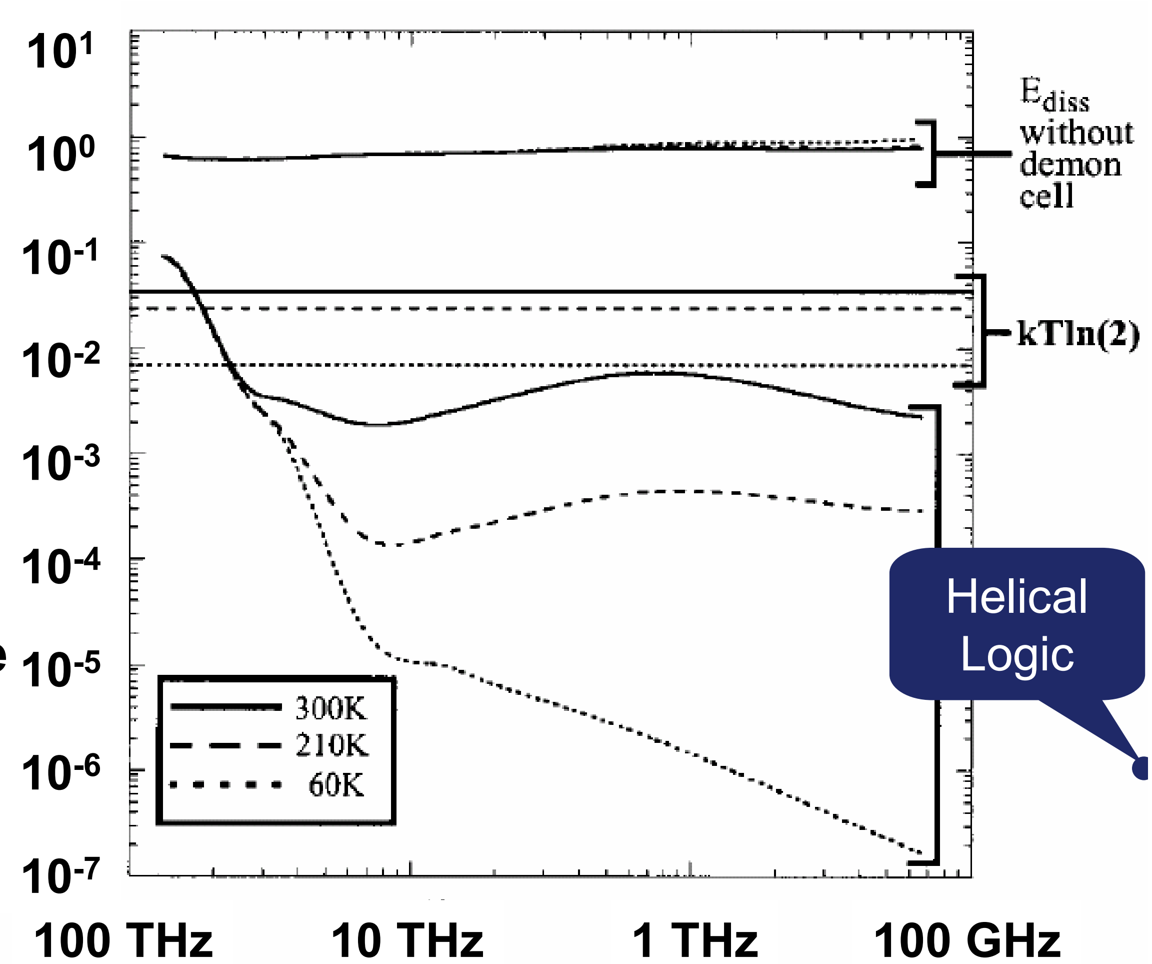




Ref. "Clocked Molecular Quantum-Dot Cellular Automata," Craig S. Lent and Beth Isaksen
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 50, NO. 9, SEPTEMBER 2003

Not Specifically Advocating Quantum Dots

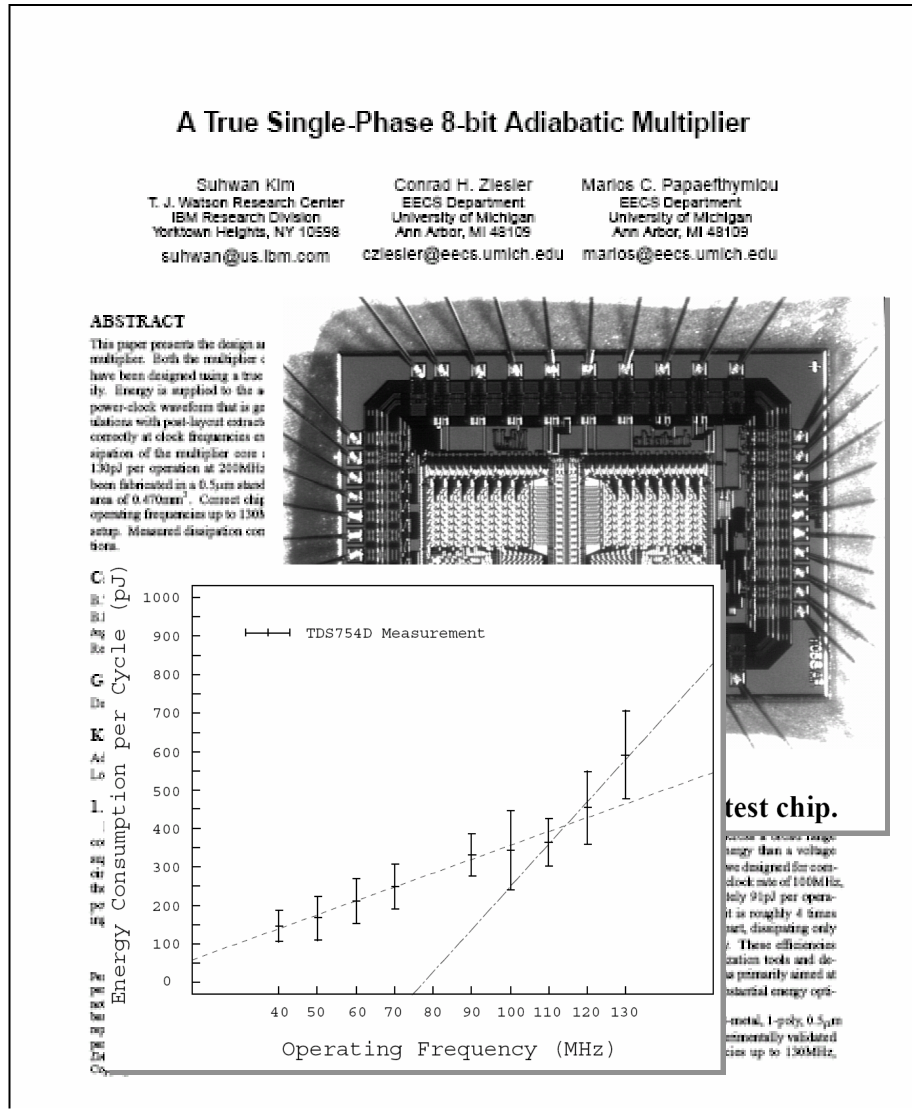
- A number of post-transistor devices have been proposed
- The shape of the performance curves have been validated by a consensus of reputable physicists
- However, validity of any data point can be questioned
- Cross-checking appropriate; see →



Ref. "Maxwell's demon and quantum-dot cellular automata," John Timler and Craig S. Lent, JOURNAL OF APPLIED PHYSICS 15 JULY 2003.
Ref. "Helical logic," Ralph C. Merkle and K. Eric Drexler, Nanotechnology 7 (1996) 325–339.

Reversible Multiplier Status

- 8×8 Multiplier Designed, Fabricated, and Tested by IBM & University of Michigan
- Power savings was up to 4:1



QCA Microprocessor Status

- M. Niemier Ph. D. Thesi
  University of Notre Dam
- 12 Bit μP
- CAD design tool princip
  - 10$\times$ circuit density of
    CMOS at same λ
- Applies to various devic
  - Metal dot 4.2 nm^2
  - Molecular 1.1 nm^2

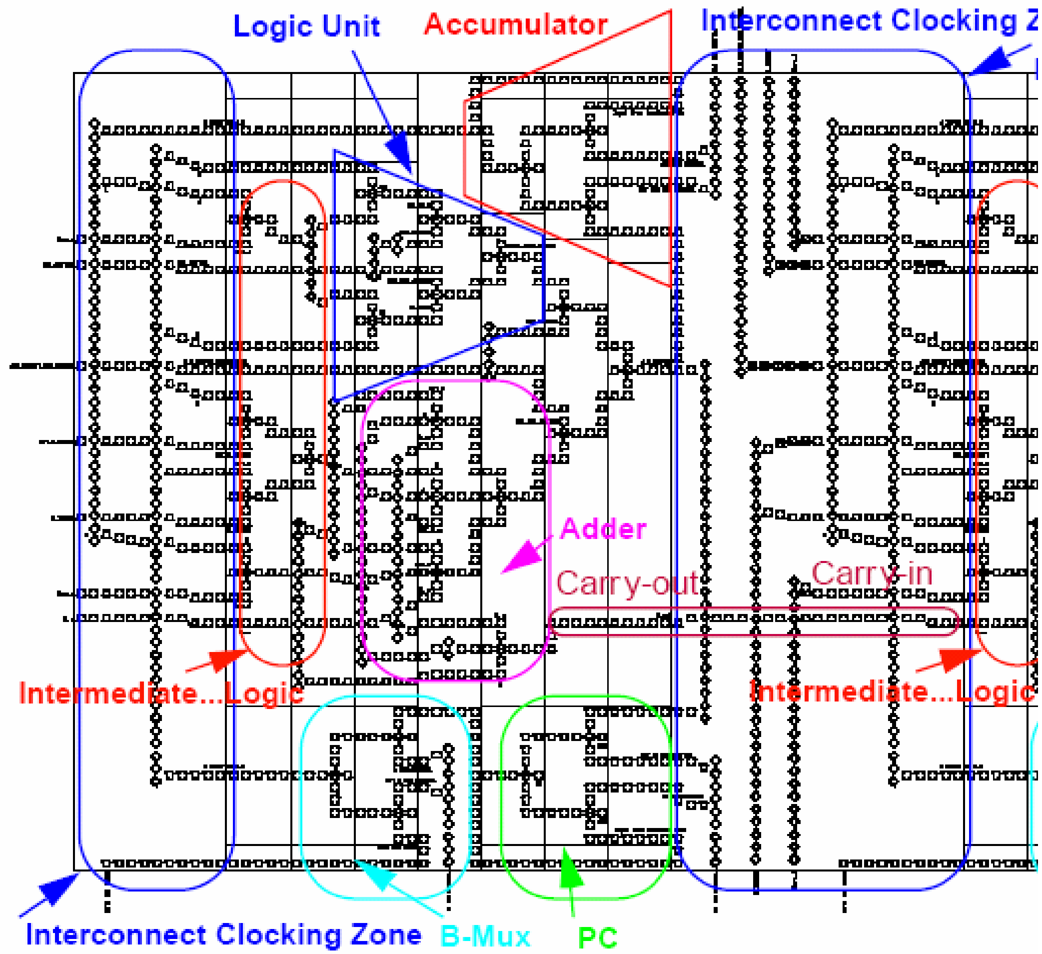
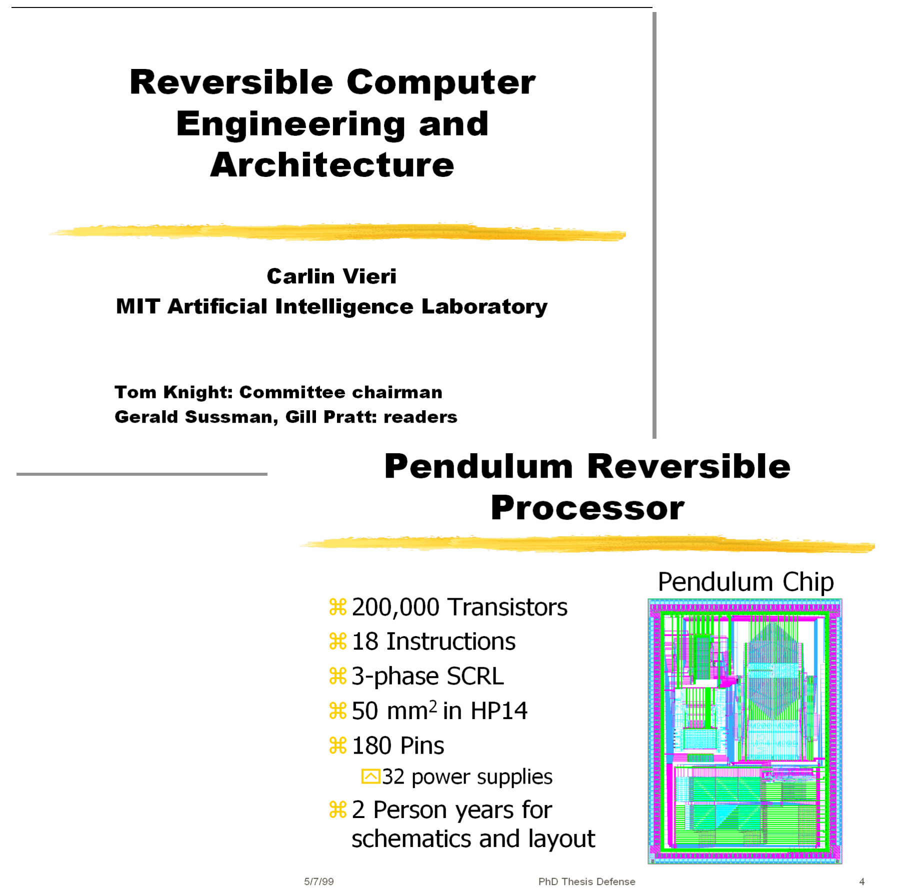


Figure 4.6. A 2-bit QCA Simple 12 ALU with registers

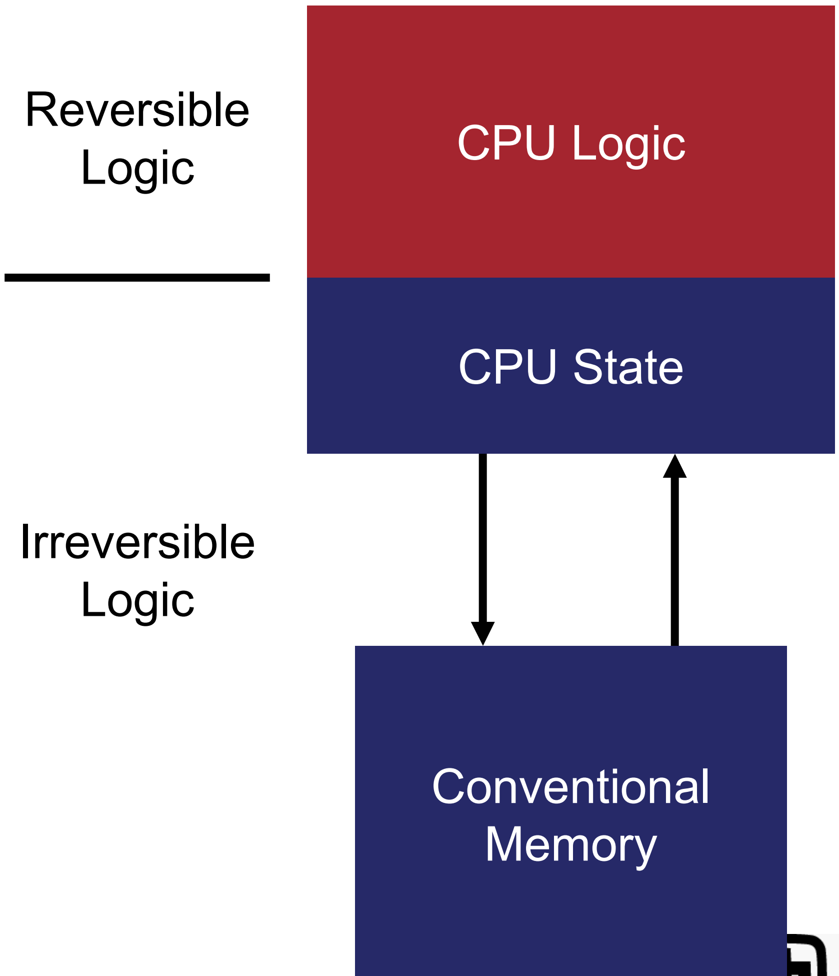
Reversible Microprocessor Status

- 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 =



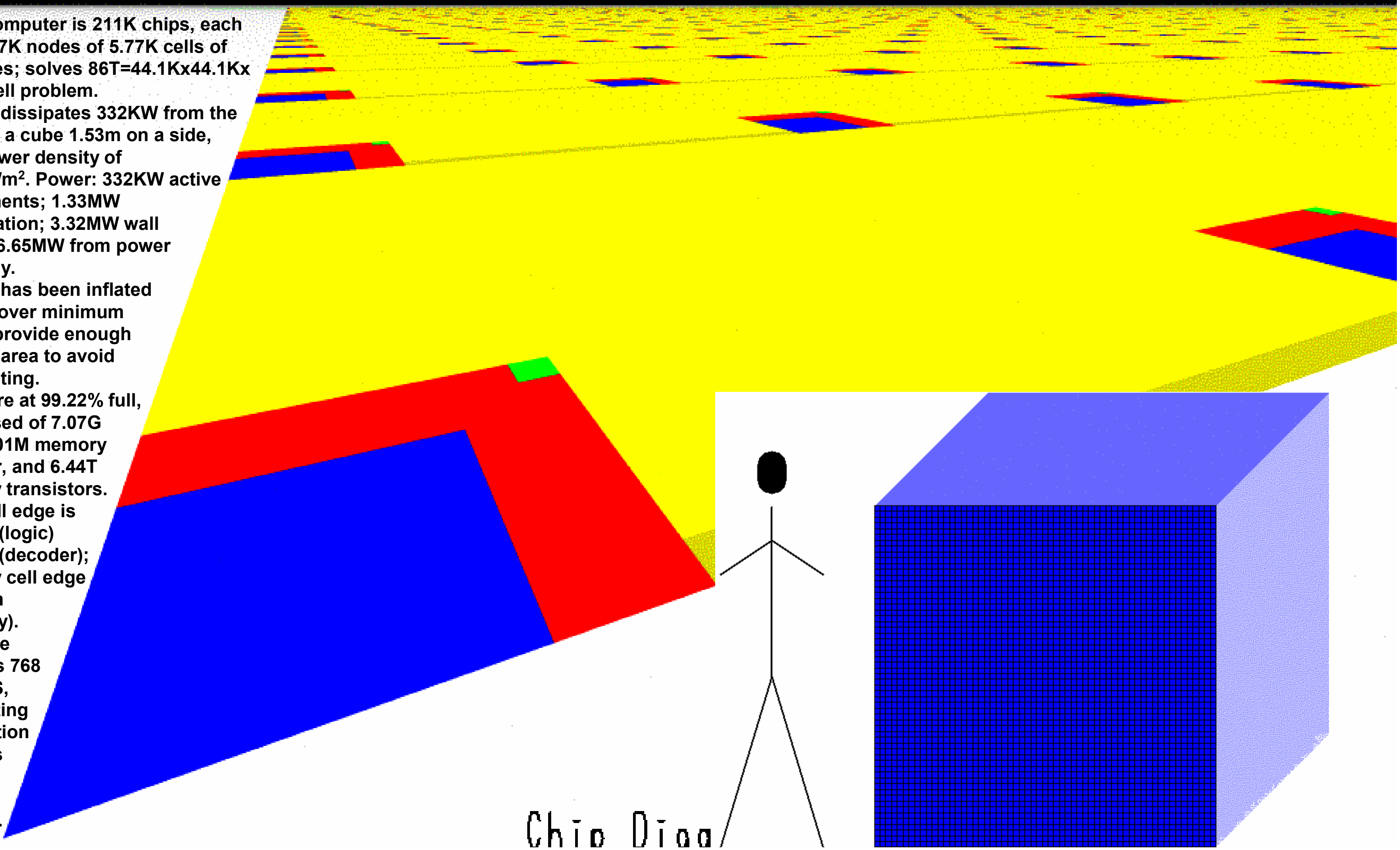
CPU Design

- 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



CTH at a Zettaflops

Supercomputer is 211K chips, each with 70.7K nodes of 5.77K cells of 240 bytes; solves 86T=44.1Kx44.1Kx 44.1K cell problem.
System dissipates 332KW from the faces of a cube 1.53m on a side, for a power density of 47.3KW/m^2. 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.



Outline

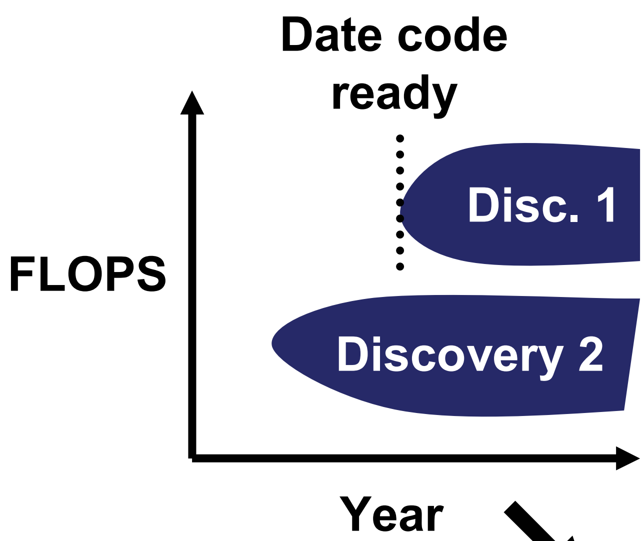
- The Computing of Physics:
  The Need for Zettaflops
- Limits of Moore's Law
  Today's Technologies
- An Expert System/Optimizer for Supercomputing
- The Physics of Computing:
  Reaching to Zettaflops
- Roadmap and Future Directions

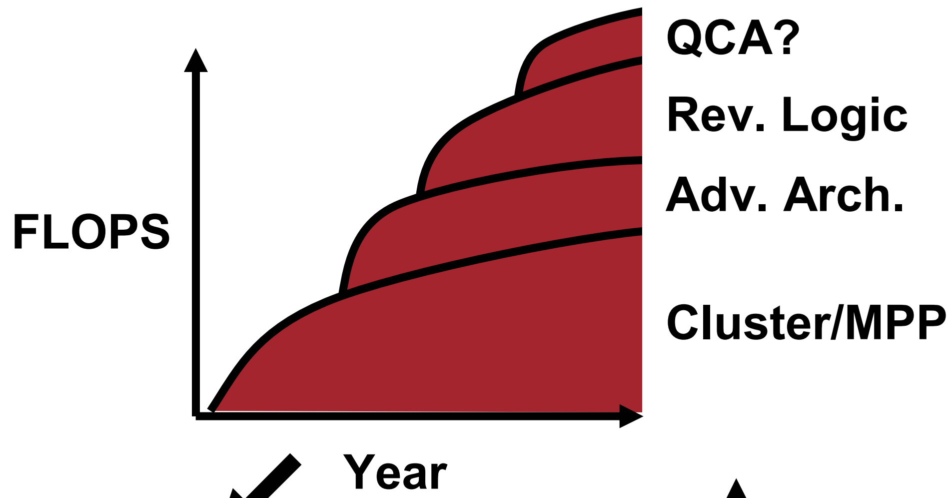
Sandia National Laboratories

Where to Go Next I: Workshop

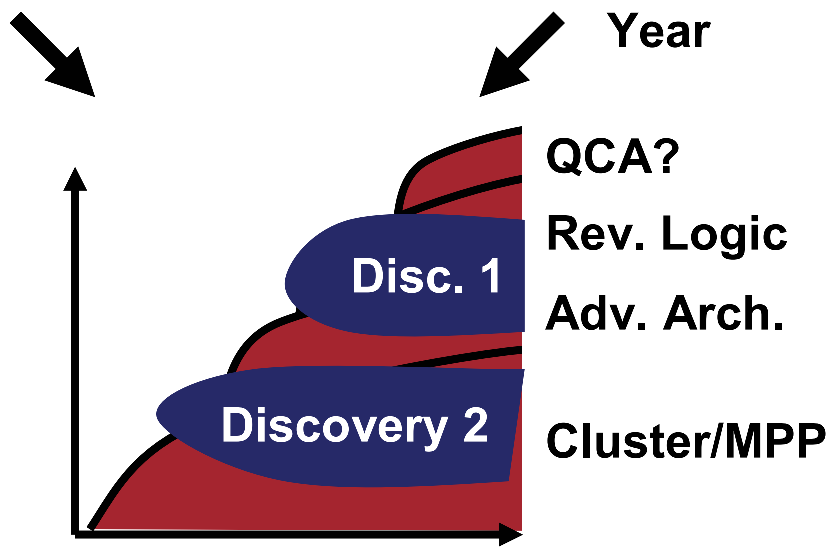
- Don’t believe me? Believe the Experts
- Workshop Agenda October 12
  - Applications session – Climate expert Phil Jones
  - Advanced Architectures – PIM expert Peter Kogge
  - Limits of Current Architectures – Me
  - Panel I: Important Applications
  - New Logic – Reversible Logic Expert Michael Frank
  - New Devices – Quantum Dot Developer Craig Lent
  - Panel II: Do Opportunities Justify the Effort?

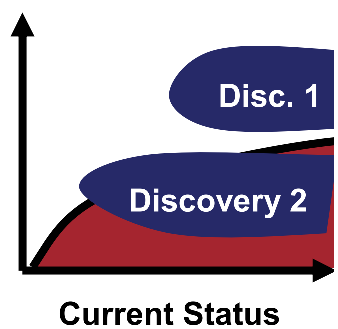
Where To Go Next II: Roadmap





Notes:
* Not necessarily one machine; different applications may require different machines
* Specifics are just my ideas





Sandia National Laboratories

Will Supercomputers Grow Forever?

- Will supercomputer simulations scale up forever, or will there be a maximum?
  - Zettaflops simulates the Earth, and the Earth is the largest thing that we care about in detail
- Will progress in science always come through "simulating physics on a computer"?
  - Perhaps future problems could be formulated as a combination of symbolic reasoning (artificial intelligence) and floating point