Review Approval

Search Detail

Submittal Details

Document Info

Title : Petaflops, Exaflops, and Zettaflops for Science and Defense

Document Number : 5231911
SAND Number : 2005-2690 C
Review Type : Electronic
Status : Approved
Sandia Contact : DEBENEDICTIS,ERIK P.
Submittal Type : Conference Paper
Requestor : DEBENEDICTIS,ERIK P.
Submit Date : 04/22/2005
Peer Reviewed? : N

Author(s)

DEBENEDICTIS,ERIK P.

Event (Conference/Journal/Book) Info

Name : Cray User Group 2005
City : Albuquerque
State : NM
Country : USA
Start Date : 05/16/2005
End Date : 05/19/2005

Partnership Info

Partnership Involved : No
Partner Approval :
Agreement Number :

Patent Info

Scientific or Technical in Content : No
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	SUMMERS,RANDALL M.	SUMMERS,RANDALL M.	04/22/2005
Conditions:			

Classification Approver	WILLIAMS,RONALD L.	WILLIAMS,RONALD L.	**04/25/2005**
Conditions:			
Manager Approver	PUNDIT,NEIL D.	PUNDIT,NEIL D.	**04/29/2005**
Conditions:			
Administrator Approver	LUCERO,ARLENE M.	KRAMER,SAMUEL	**06/05/2007**

SAND2005-2690C

Petaflops, Exaflops, and Zettaflops for Science and Defense

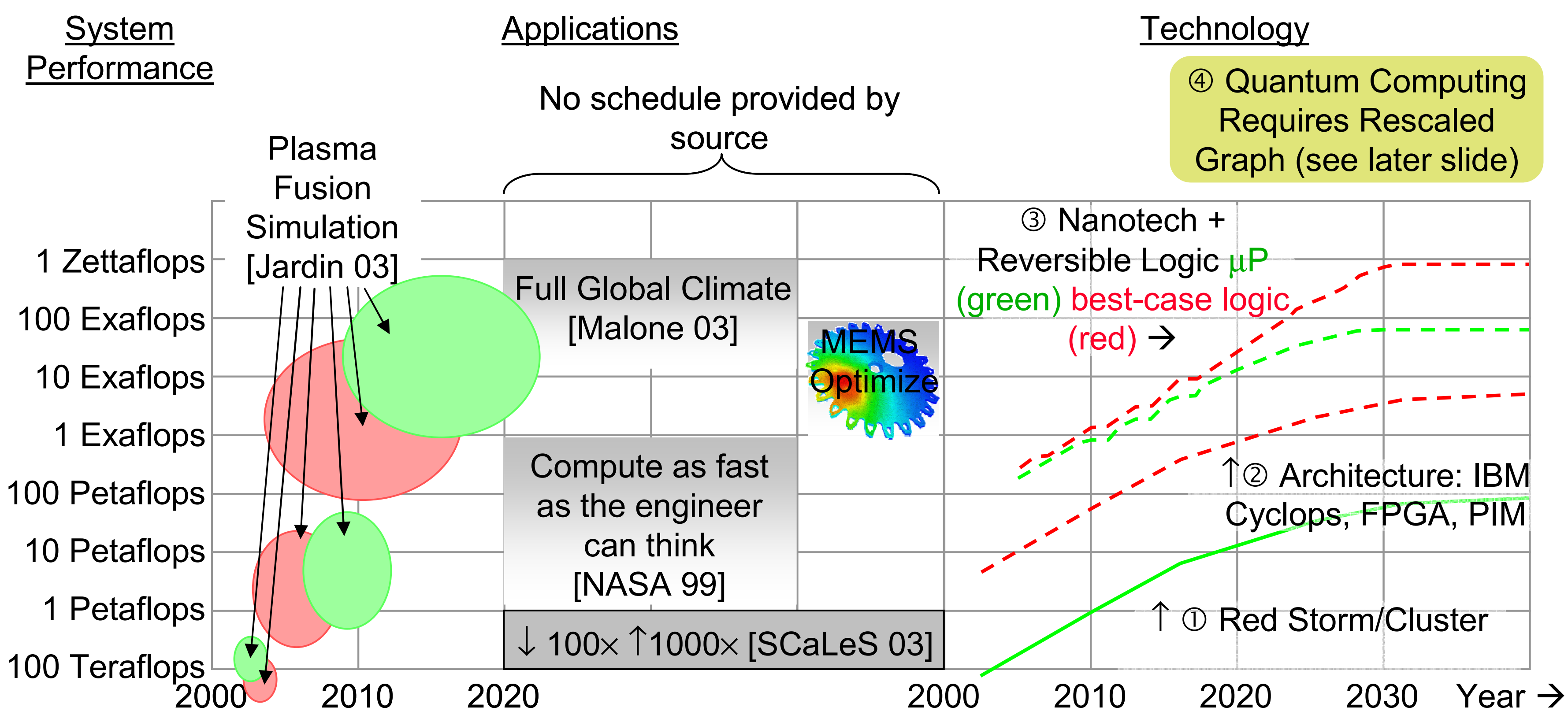
Erik P. DeBenedictis
Sandia National Laboratories

May 16, 2005

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

Applications and $100M Supercomputers



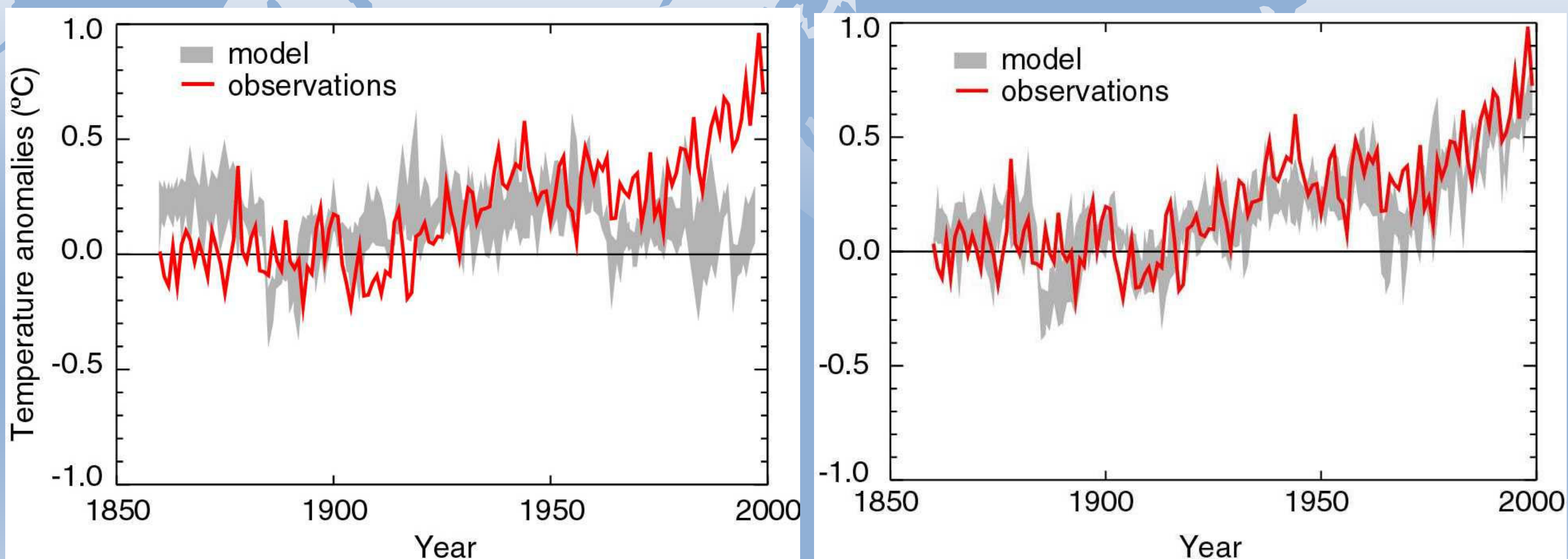
[Jardin 03] S.C. Jardin, "Plasma Science Contribution to the SCaLeS Report," Princeton Plasma Physics Laboratory, PPPL-3879 UC-70, available on Internet.
[Malone 03] Robert C. Malone, John B. Drake, Philip W. Jones, Douglas A. Rotman, "High-End Computing in Climate Modeling," contribution to SCaLeS report.
[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.
[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.

Outline

- **Exemplary Zettaflops Problems**
- **The Limits of Moore's Law**
- **Beyond Moore's Law**
  - **Industry's Plans**
  - **Nanotech and Reversible Logic**
  - **Quantum Computing**
- **Conclusions**

Sandia National Laboratories

Simulation of Global Climate



Stott et al, Science 2000

“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

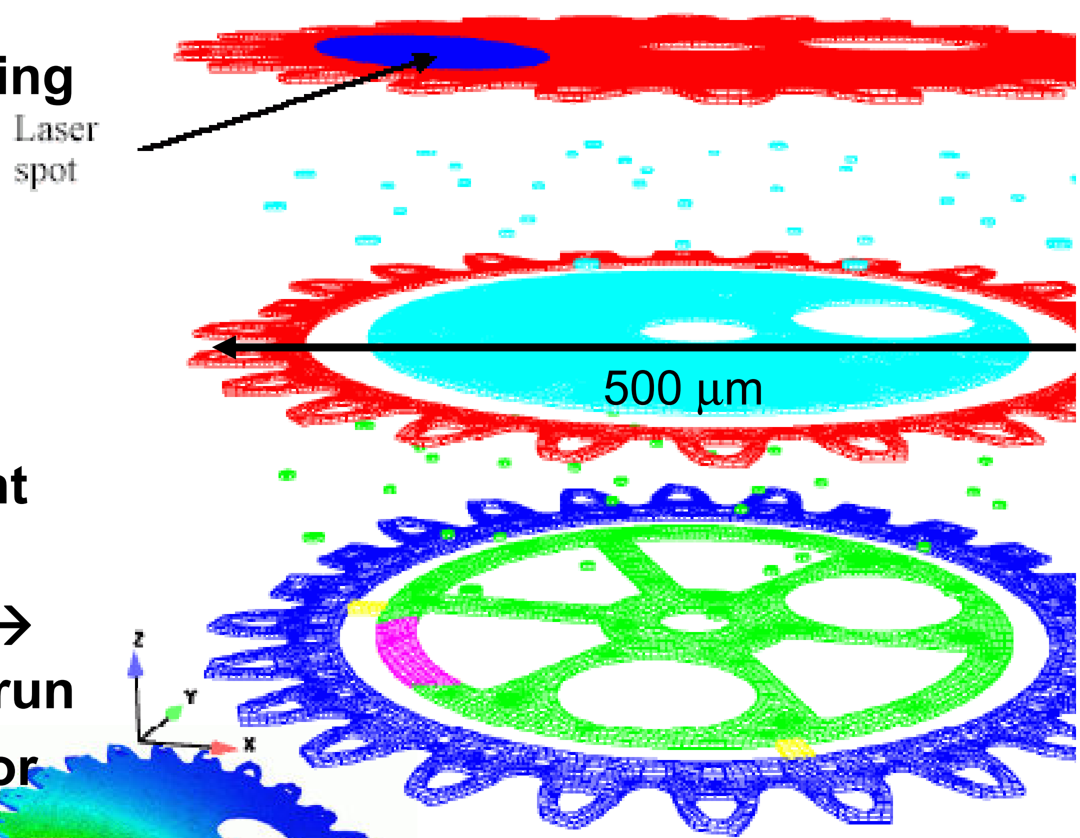
- 1 Zettaflops
- 100 Exaflops
- 1 Exaflops
- 10 Petaflops
- 100 Teraflops
- 10 Gigaflops

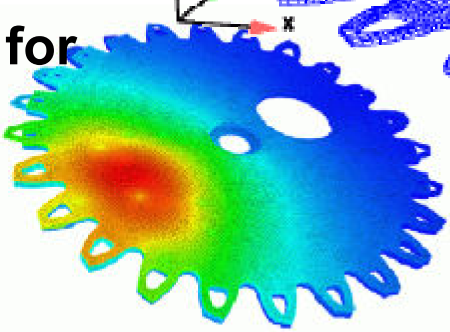
Issue	Scaling
Ensembles, scenarios 10×	Embarrassingly Parallel
Run length 100×	Longer Running Time
New parameterizations 100×	More Complex Physics
Model Completeness 100×	More Complex Physics
Spatial Resolution $10^4\times$ ($10^3\times$-$10^5\times$)	Resolution
Clusters Now In Use (100 nodes, 5% efficient)	

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)

Exemplary Exa- and Zetta-Scale Simulations

- Sandia MESA facility using MEMS for weapons
- Heat flow in MEMS not diffusion; use DSMC for phonons
- Shutter needs 10 → Exaflops on an overnight run for steady state
- Geometry optimization → 100 Exaflops overnight run
  - Adjust spoke width for high b/w no melting





Sandia National Laboratories

FLOPS Increases for MEMS

- 100 Exaflops
- 10 Exaflops
- 30 Petaflops
- 600 Gigaflops
- 5 Gigaflops

Issue	Scaling
Optimize 10×	Sequential
Run length 300×	Longer Running Time
Scale to 500μm²×12μm disk 50,000×	Size
2D → 3D 120×	Size
2μm×.5μm×3μs 2D film 10 × 1.2 GHz PIII	

Sandia National Laboratories

Outline

- Exemplary Zettaflops Problems
- The Limits of Moore's Law
- Beyond Moore's Law
  - Industry's Plans
  - Nanotech and Reversible Logic
  - Quantum Computing
- Conclusions

*** 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 (← 125:1 →)	Derate for manufacturing margin (4×)	Estimate
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

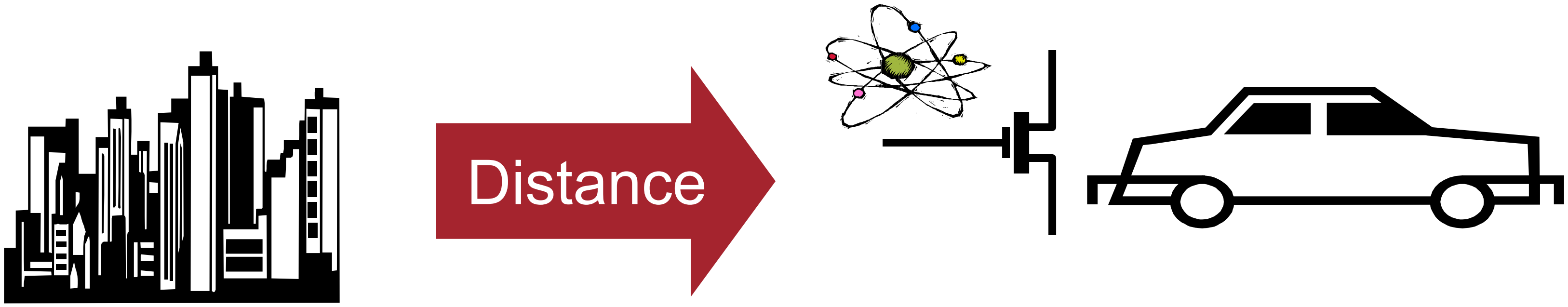
Metaphor: FM Radio on Trip to Santa Fe

- You drive to Santa Fe 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

FM Radio and End of Moore’s Law



Driving away from FM transmitter→less signal
Noise from electrons → no change

Increasing numbers of gates→less signal power
Noise from electrons → no change

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.

Scientific Supercomputer Limits

Best-Case Logic	Microprocessor Architecture		
2×10^{24} logic ops/s			
Expert Opinion: 100 Exaflops	800 Petaflops		
← 125:1 →			
Estimate: 25 Exaflops	200 Petaflops		
4 Exaflops	32 Petaflops		
1 Exaflops	8 Petaflops		
	80 Teraflops		
	40 Teraflops		

Physical Factor	Source of Authority
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)
Derate for manufacturing margin (4×)	Estimate
Uncertainty (6×)	Gap in chart
Improved devices (4×)	Estimate
Projected ITRS improvement to 22 nm (100×)	ITRS committee of experts
Lower supply voltage (2×)	ITRS committee of experts
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

Outline

- Exemplary Zettaflops Problems
- The Limits of Moore's Law
- Beyond Moore's Law
  - Industry's Plans
  - Nanotech and Reversible Logic
  - Quantum Computing
- Conclusions

Proceeding

- So industry has plans to extend Moore's Law, right?
  - Next slide shows ITRS Emerging Research Devices (ERD), the devices under consideration by industry
  - All are either hotter, bigger, or slower
  - Erik is now on ITRS ERD committee
- What is scientifically feasible for Gov't funding?
  - Nanotechnology
    - Efforts all over
  - Reversible logic
    - Odd name for a method of cutting power below k_BT
    - Not currently embraced by industry
  - Quantum computing
    - More later

Sandia National Laboratories

ITRS Device Review 2016

Technology	Speed (min-max)	Dimension (min-max)	Energy per gate-op	Comparison
CMOS	30 ps-1 μs	8 nm-5 μm	4 aJ	
RSFQ	1 ps-50 ps	300 nm- 1μm	2 aJ	Larger
Molecular	10 ns-1 ms	1 nm- 5 nm	10 zJ	Slower
Plastic	100 μs-1 ms	100 μm-1 mm	4 aJ	Larger+Slower
Optical	100 as-1 ps	200 nm-2 μm	1 pJ	Larger+Hotter
NEMS	100 ns-1 ms	10-100 nm	1 zJ	Slower+Larger
Biological	100 fs-100 μs	6-50 μm	.3 yJ	Slower+Larger
Quantum	100 as-1 fs	10-100 nm	1 zJ	Larger

Data from ITRS ERD Section.

Outline

- Exemplary Zettaflops Problems
- The Limits of Moore's Law
- Beyond Moore's Law
  - Industry's Plans
  - Nanotech and Reversible Logic
  - Quantum Computing
- Conclusions

The Parallelism Issue

- **Initially, didn't meet constraints**

More Parallelism

Scaled Climate Model

2D → 3D mesh,
one cell per processor

Parallelize cloud-resolving model and ensembles

One Barely Plausible Solution

Consider special purpose logic with fast logic and low-power memory

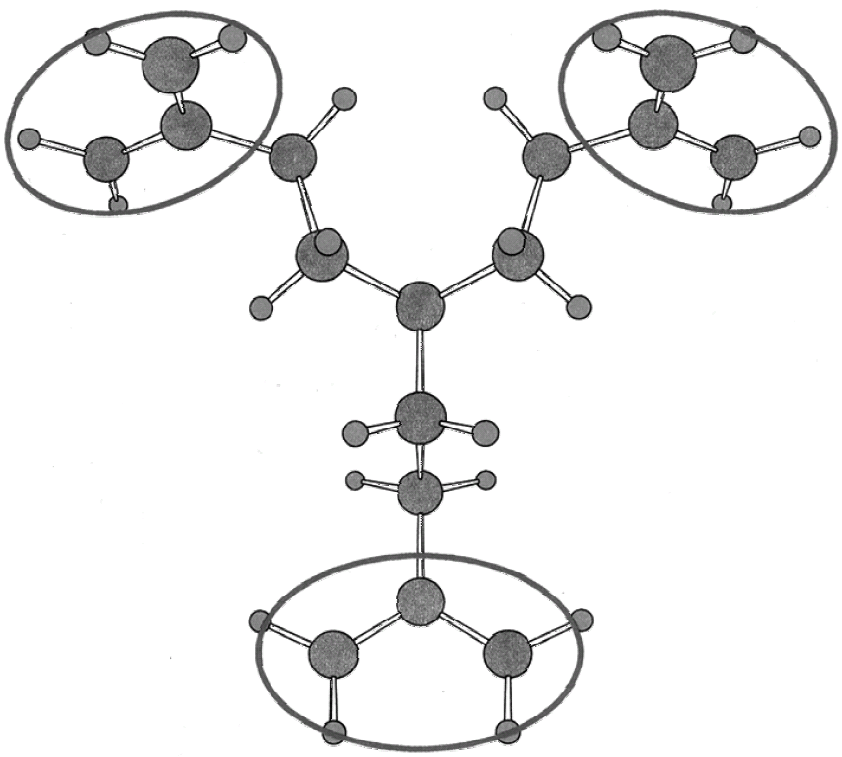
Consider only highest performance published nanotech device QDCA

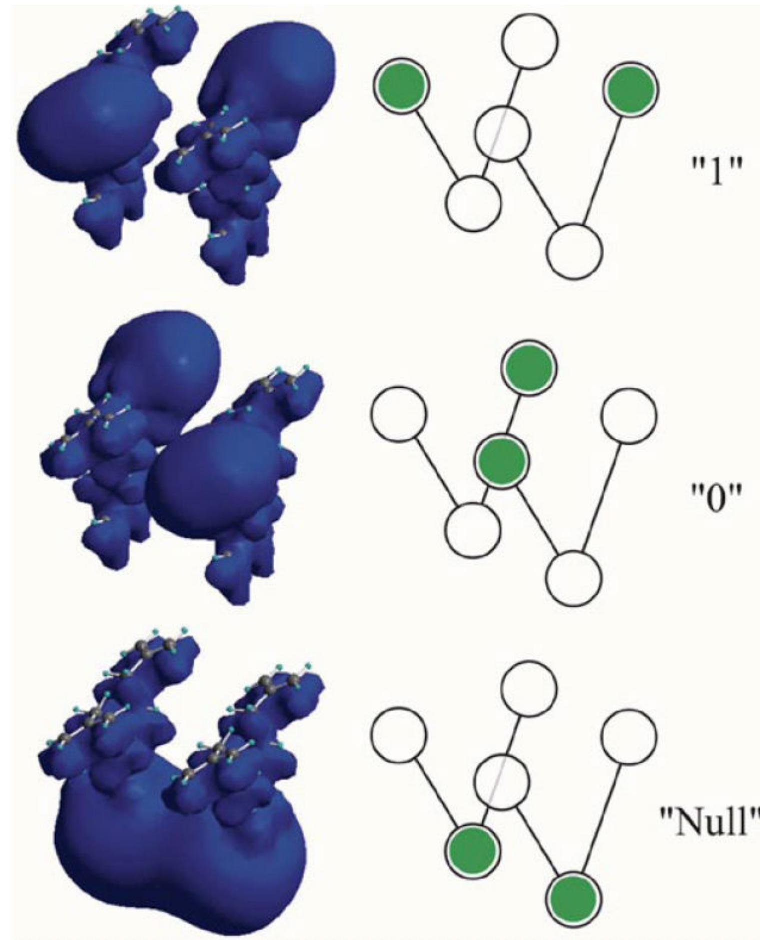
Initial reversible nanotech

More Device Speed

An Exemplary Device: Quantum Dots

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

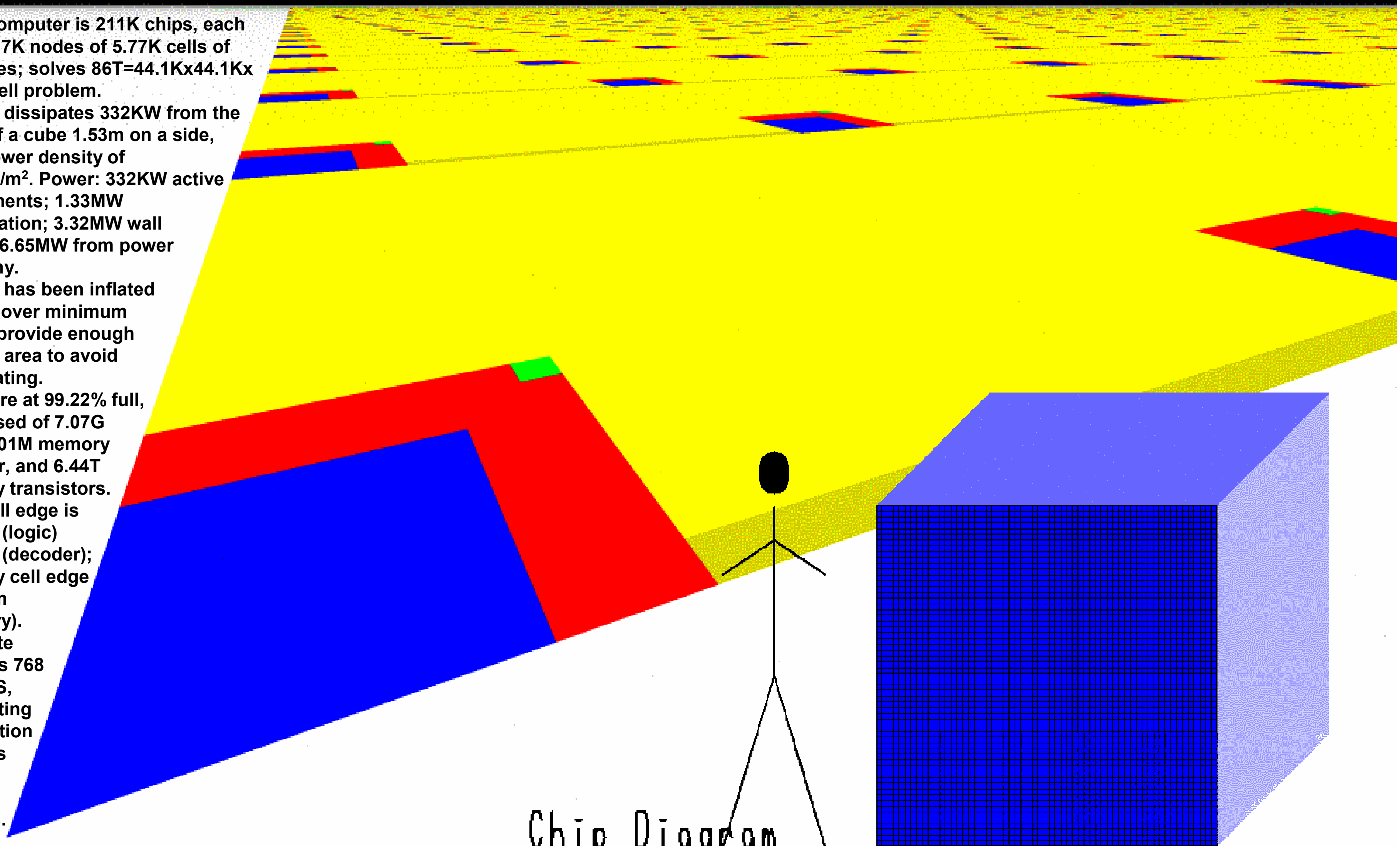




Ref. "Clocked Molecular Quantum-Dot Cellular Automata," Craig S. Lent and Beth Isaksen
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 50, NO. 9, SEPTEMBER 2003

Atmosphere Simulation at a Zettaflops

Supercomputer is 211K chips, each with 70.7K nodes of 5.77K cells of 240 bytes; solves 86T=44.1Kx44.1Kx 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.

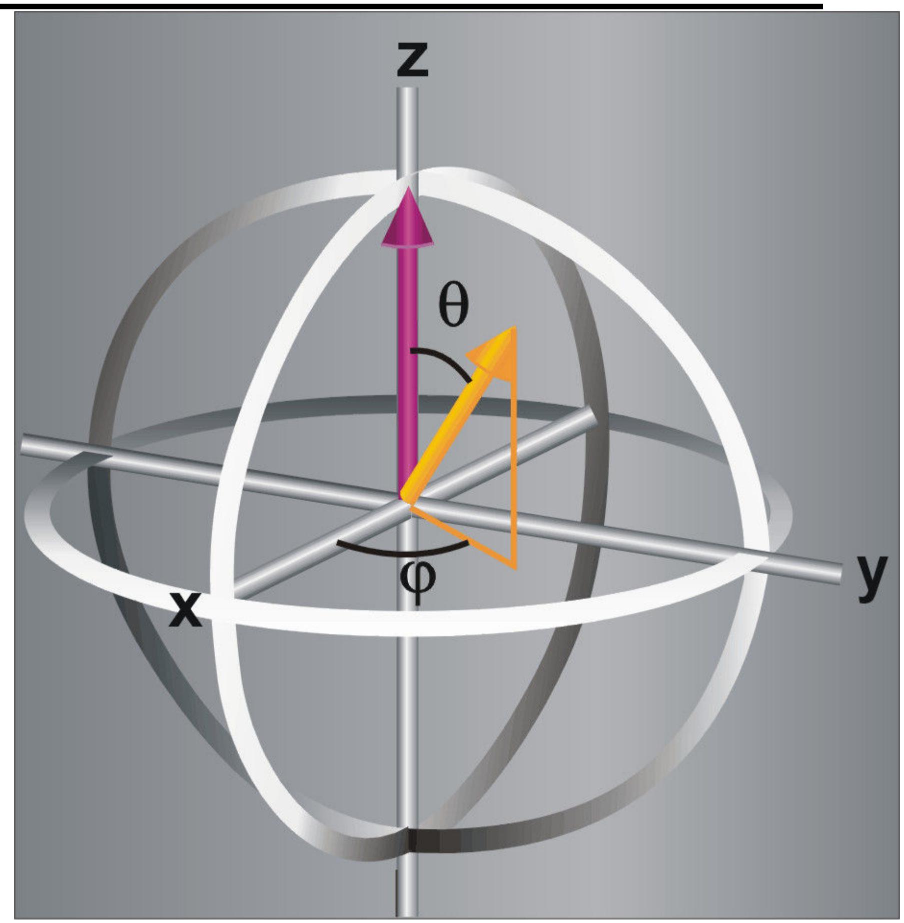


Outline

- Exemplary Zettaflops Problems
- The Limits of Moore's Law
- Beyond Moore's Law
  - Industry's Plans
  - Nanotech and Reversible Logic
  - Quantum Computing
- Conclusions

Why Quantum Computing is Interesting

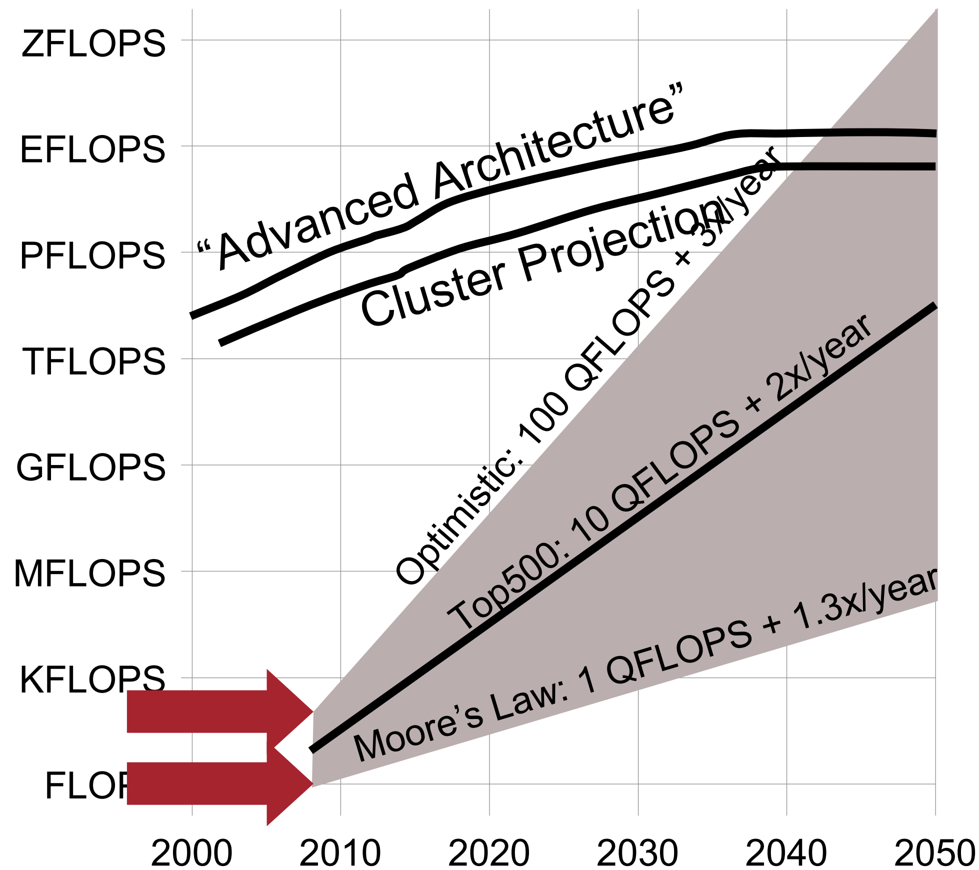
- A Superset of Digital
  - Spin “up” is a 1
  - Spin “down” is a 0
  - Other spins
    - Sidewise
    - Entangled
    - Phase
  - Like wildcards
    - 1011??????
    - Up to 2^N states → in “quantum parallel”



Sandia National Laboratories

Emergence of Quantum Computing

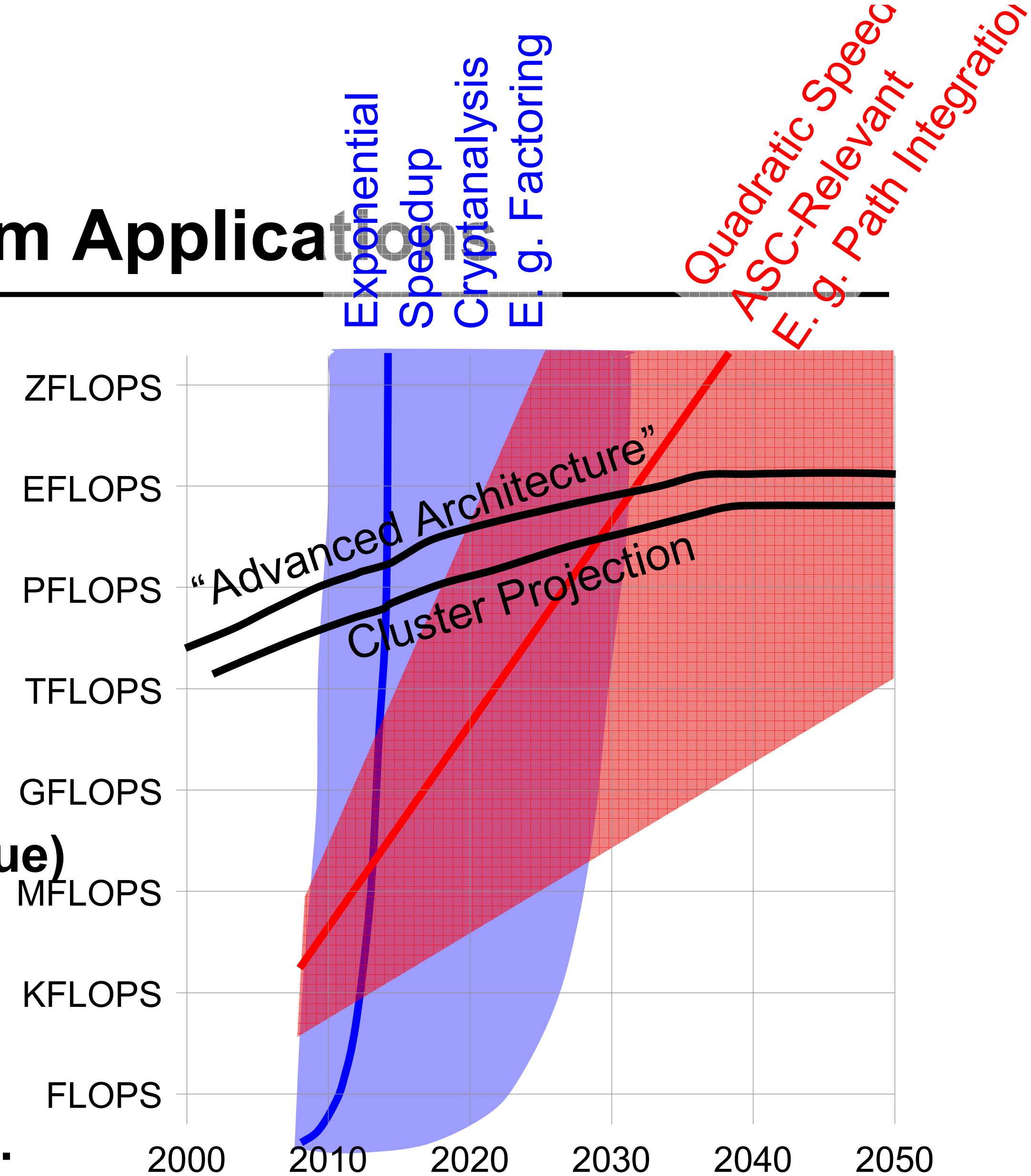
- There appears to be an engineering case for quantum computers of 1-100 Q-FLOPS
- One would expect an exponential growth rate for quantum computers similar to Moore's Law, but the rate constant is impossible to predict, so three possibilities have been graphed



Ref. "How to build a 300 bit, 1 Gop quantum computer," Andrew M. Steane, Clarendon Laboratory, UK, quant-ph/0412165

Quantum Applications

- Consider the classical computer equivalent to a Quantum Computer
- First use believed to be factoring in crypt-analysis, with expo-nential speedup over classical computers (blue)
- Second, a quantum computer can also be used for other applications (pink) with quadratic speedup (e. g. Actinide chemistry)



Sandia National Laboratories

Outline

- Exemplary Zettaflops Problems
- The Limits of Moore's Law
- Beyond Moore's Law
  - Industry's Plans
  - Nanotech and Reversible Logic
  - Quantum Computing
- Conclusions

Conclusions

- Important applications exist to 1 Zettaflops
- Performance of $100M μP-based supercomputers will rise to only ~30-200 Petaflops
  - This will be sufficient to meet all existing plans
  - However, there are many apparently valid uses of computers that exceed these limits, but where there is no commitment at this time
- Advanced Architectures (e. g. PIM) will rise to ~4-25 Exaflops
  - Cray Cascade moves in this direction
- Nanotech and Reversible logic good to perhaps 1 Zettaflops
- Quantum computing
  - Will not help existing code
  - Blasts out of the top of the chart for new codes