SF 1008-RA (FRONT 9-2003)
Supersedes (8-2003) issue

Information Release

SAND 2004-0680P

REVIEW & APPROVAL FORM

Originating organization: Please complete Sections 1 - 6. Print or type all information. See attached instruction sheets for additional information.

This form is used to review and approve information releases before they are released outside of Sandia.

Public releases must go through the Formal R&A Process, in which case this form must be completed through Section 10.
For releases going through the Organizational R&A Process, organizational management is encouraged to complete this form through Section 6 and to file it for future reference.
This form can also be used (through Section 6 & Section 8) to document approvals when an information release changes distribution limitations (e.g., when Internal Distribution Only becomes Unlimited Release).
For information on which R&A process to use, or additional R&A information:

- See the "Review and Approval for Communication" webpage: http://www-irn.sandia.gov/recordsmgmt/revapprov/revapprov.htm or
- Contact Linda Cusimano (NM), (505) 844-4980 ; Kelly McClelland (CA), (925) 294-2311

SECTION 1. Protecting Sandia and Partnership Interests.

Is this release the result of ☐ a CRADA? ☐ Work for Others? ☐ Other partnership, MOU agreement, funding source, or understanding OF ANY KIND? ☐ Information controlled by other agencies

☒ If NO: Go to Section 2.
☐ If YES: Agreement Number is

Has your partner, non-SNL information owner, or funding agency given approval for this release? ☐ Yes ☐ No

SECTION 2. Document Title and Author Information:

Full title of document **Issues In Supercomputer Architecture For The Next Dozen Years**

Author or contact (Sandian) **Erik P. DeBenedictis** (Full First Name Full Middle Name Full Last Name) Signature

Phone No. **505 284 4017** E-Mail Address **epdeben@sandia.gov** Org. No. **9223** Mail Stop No. **1110**

Project number **7101** Task number **13.04** (Identifies funding source – will not be charged)

☐ Contract Author to Sandia. (Contractor's name and contract no.)

SECTION 3. Document Format and Release Event Information. Indicate the planned format(s) of the information release, as well as information about the release event.

Document Format(s): ☐ SAND Report ☐ Abstract ☐ Conference Paper ☐ Journal Article ☐ Sandia Open Network (External) ☐ Computer Software ☒ Publication (all other types of publications including reports, vugraphs, posters, exhibits, displays, videos, brochures, internal memoranda, newsletters, factsheets)

Release Event: Indicate the name of the conference, meeting, or publication, the sponsoring organization, and the place and date of event. If this release is an electronic posting, provide the current viewing address and intended posting location.

Name of Conference / Journal / Book: **Presentation at New Mexico State University**
Place of Event: **New Mexico State University** Las Cruces, NM Date: **1/29/04 thru 1/29/04**
Internet Address of Electronic Posting:

SECTION 4. Classification and Sensitivity of Information. Contact Classification Dept. 4225 (8511 - CA) for questions.

Indicate classification level and category of information release or whether information release is unclassified:
Classification of: Document Title U Document Abstract N/A The Document U

☐ Classified - Limited Release. Indicate additional access restrictions:
☐ NWD Sigma ☐ CNWDI ☐ NOFORN ☐ Program Designated Special Handling (Distribution Limitation) ☐ Other

☐ Unclassified - Limited Release. Indicate all Unclassified Controlled Information (UCI) categories access restrictions:
☐ Export Controlled Information (ECI) ITAR/EAR/
☐ Non-Sandia Proprietary Information
☐ Official Use Only (OUO) Exemption No.
☐ Patent Caution
☐ Protected CRADA Information (Release date)
☐ Program Designated Special Handling (Distribution Limitation)
☐ Unclassified Controlled Nuclear Information (UCNI)
☐ Other (specify)

☒ Unclassified - Unlimited Release. Information is unclassified with no access restrictions, i.e., distribution may be made worldwide.

☐ **DUSA Exemption -** This information is released under DUSA Exemption (Mark appropriate sensitivity above. **Section 8 review not required.**)

Derivative Classifier (DC) who is knowledgeable of information sensitivity or DUSA Delegate: Paul Yarrington (Name) Paul Yarrington (Signature) 9233 (Org.) 3/5/04 (Date)

(Continued on page 2)

SF 1008-RA (9-2003)
Supersedes (8-2003) issue

SECTION 5. Disclosure of Technical Advance

A Technical Advance is an original achievement or non-obvious progress in a scientific or engineering sense, including the creation of soft
It may be protected by patent or copyright. The Originators of a Technical Advance may be inventors or authors.

Does the subject of this Information Release represent a Technical Advance as defined above?
☐ Yes ☒ No If **No**, go to Section 6.

If **Yes**, has a Disclosure of **Technical Advance (TA)**, Form SF 1155-TD, been filed with the Sandia Patent and Licensing Center?
☐ Yes SD No. _____ ☐ No If No, please follow up with a TA form obtainable from:

SECTION 6. Line/Program Signatures and Approvals. Print or type all author information; obtain appropriate signatures from next-level manager. Where concurrence is obtained in case of multiple authors, approval need only go through the principal author's line organization.

Authors' Names (Print or type) (Full First, Middle, & Last Name)	Org. No./ Mail Stop	Phone No.	Next Level Manager's Signature	Date
Erik P. DeBenedictis (Full First Name Full Middle Name Full Last Name)	**1110**	**505 284 4017**	Neil Pundit	**2/3/04**
(Full First Name Full Middle Name Full Last Name)				
(Full First Name Full Middle Name Full Last Name)				
(Full First Name Full Middle Name Full Last Name)				
(Full First Name Full Middle Name Full Last Name)				

Program Manager's Name and Signature _____

THE FOLLOWING SECTIONS ARE USED FOR THE FORMAL REVIEW & APPROVAL PROCESS.

SECTION 7. To be completed by the Patent and Licensing Department (NM: 11500/MS 0161, CA: 11600/MS 9031).

Copyright Interest?	☐ Yes	☒ No	If Yes, copyright may be asserted, subject to DOE approval.
Patent Interest?	☐ Yes	☒ No	If Yes, TA form has been or should be submitted.
Patent Caution?	☐ Yes	☒ No	If Yes, TA form has been or should be submitted and dissemination will be limited.

Patent Attorney's/Agent's Signature M.A.O. Peterson, Date 02/11/04

SECTION 8. To be completed by the Classification and Sensitive Information Department (NM: 4225/MS 0175, CA 8511, MS 9021).

Signature Ronald Wilkerson Date 2/17/04

SECTION 9. To be completed by the Promotional Communications Departments (NM: 9612/MS 0612, CA: 8815/MS 9021).

Promotional Communications must be reviewed for adherence to Corporate "Common Look and Feel" guidelines by PR & Communications Center 12600, MS 0619 (NM) or by Communications & Public Affairs Dept. 8528, MS 9131 (CA). In addition, all publications (including SAND Reports and fact sheets) distributed outside of Sandia that use color (ink) **printing**, as well as all internal products of two or more colors that use color **printing**, must go through a Section 9 review. **NOTE**: SAND Reports and internal release products that use color **copying** do **not** need a Section 9 review. (Contact Printing & Duplicating Dept. 12630 to get a head start on DOE approval.)

DOE approval received _____ Date Signature _____ Date _____

SECTION 10. To be completed by the Review & Approval Desk (NM: 9612/MS 0612, CA: 8511/MS 9021).

Approved under the following conditions:

1. **That the following statement is printed on the document: 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.**
2. **That, if the release has been determined to be sensitive or classified, all the appropriate markings are on the released item;**
3. **That all edits requested by reviewers are completed before release;**
4. **That the final version is submitted to the Technical Library.**

Signature Dorothy Martin Date FEB 19 2004

SAND2004-0680P

Issues In Supercomputer Architecture For The Next Dozen Years

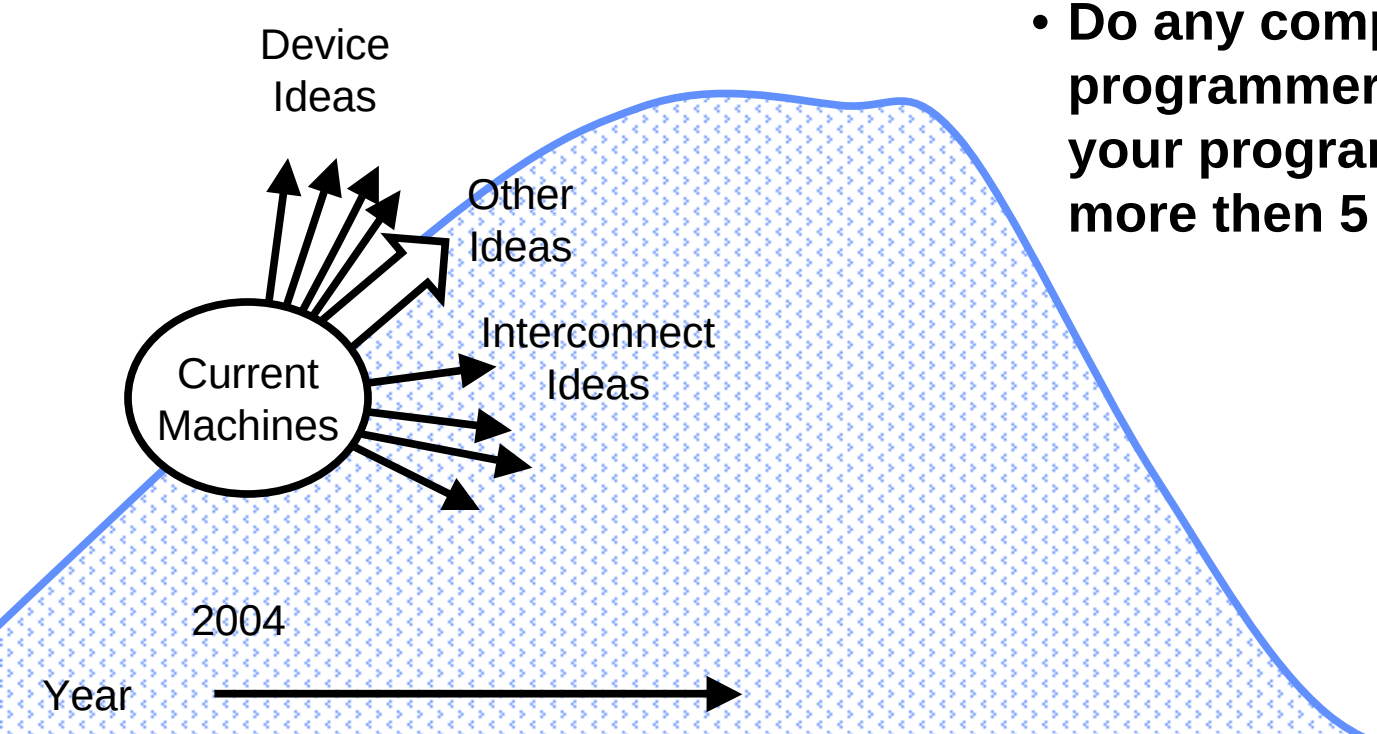
Erik P. DeBenedictis

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

Sandia National Laboratories

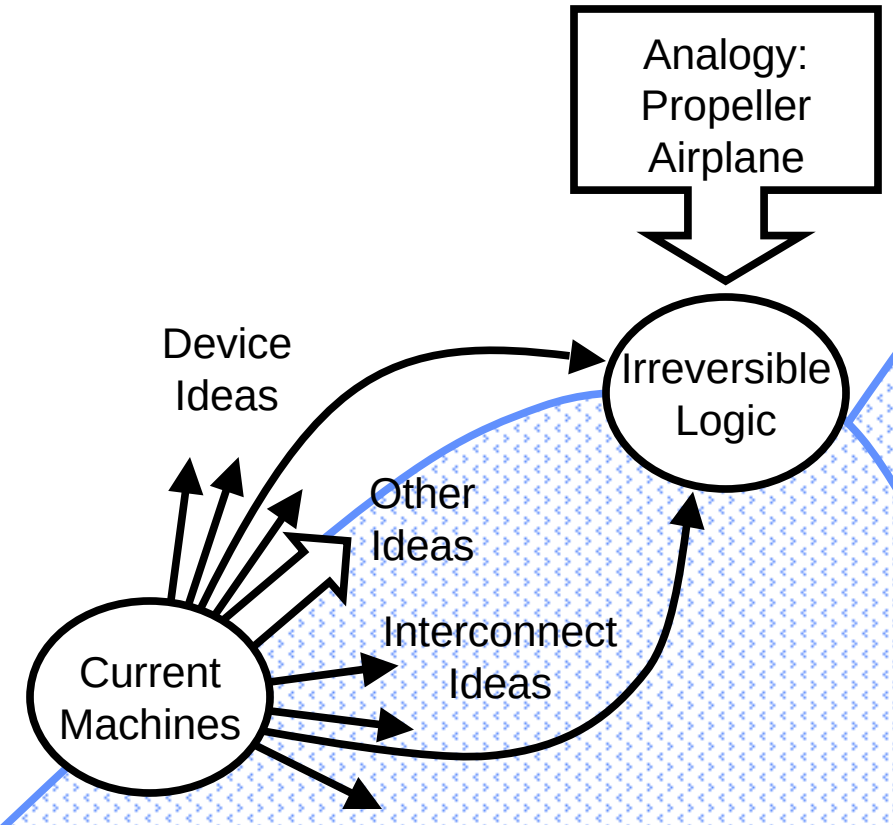
Programmatics

- **Currently, we are heading up the mountain of "Moore's Law"**



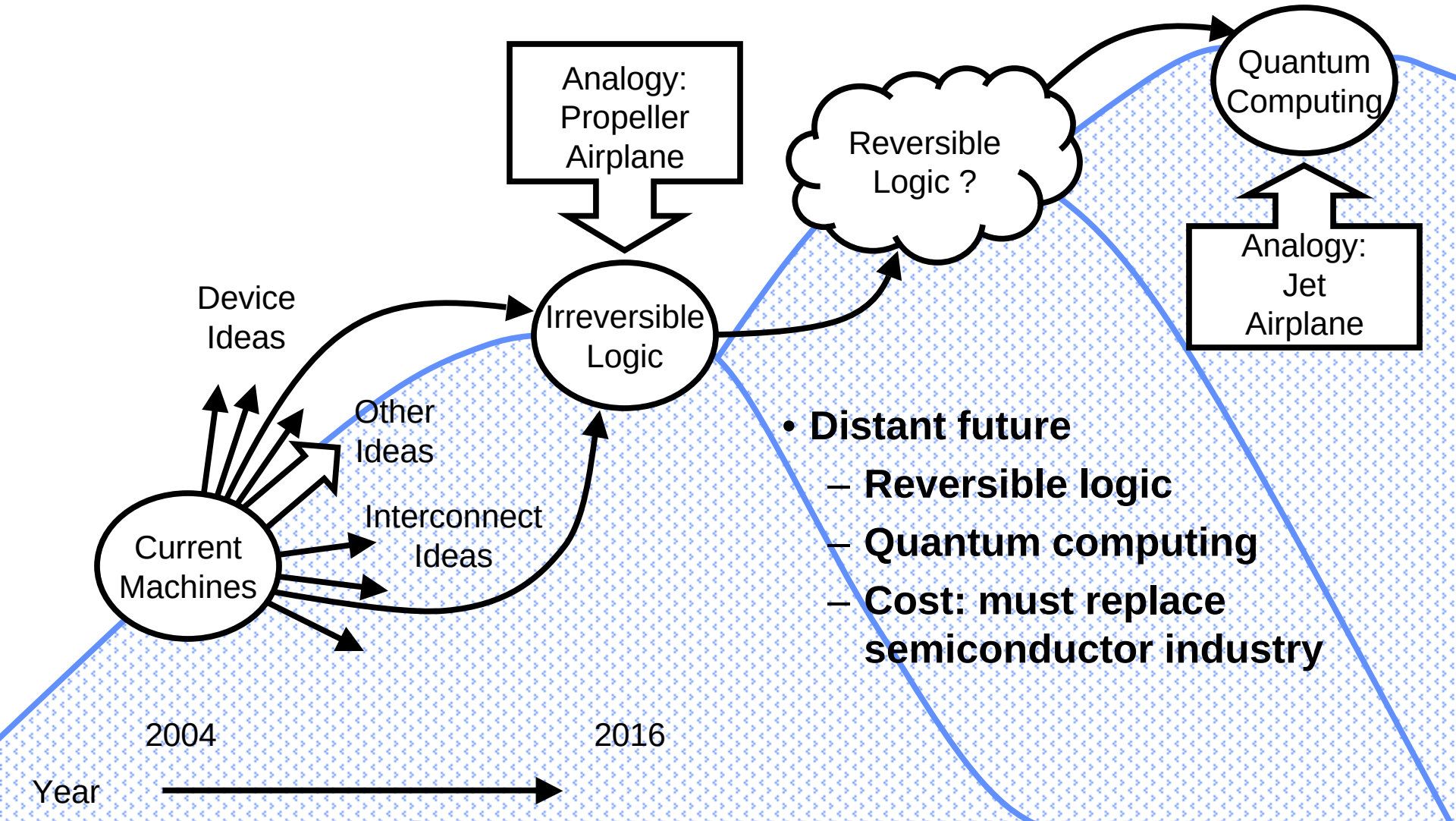
- **An architect has an idea of what their building will look like in 10, 20, 50, 100 years**
- **Do any computer programmers know how your programs will run in more then 5 years?**

Programmatics

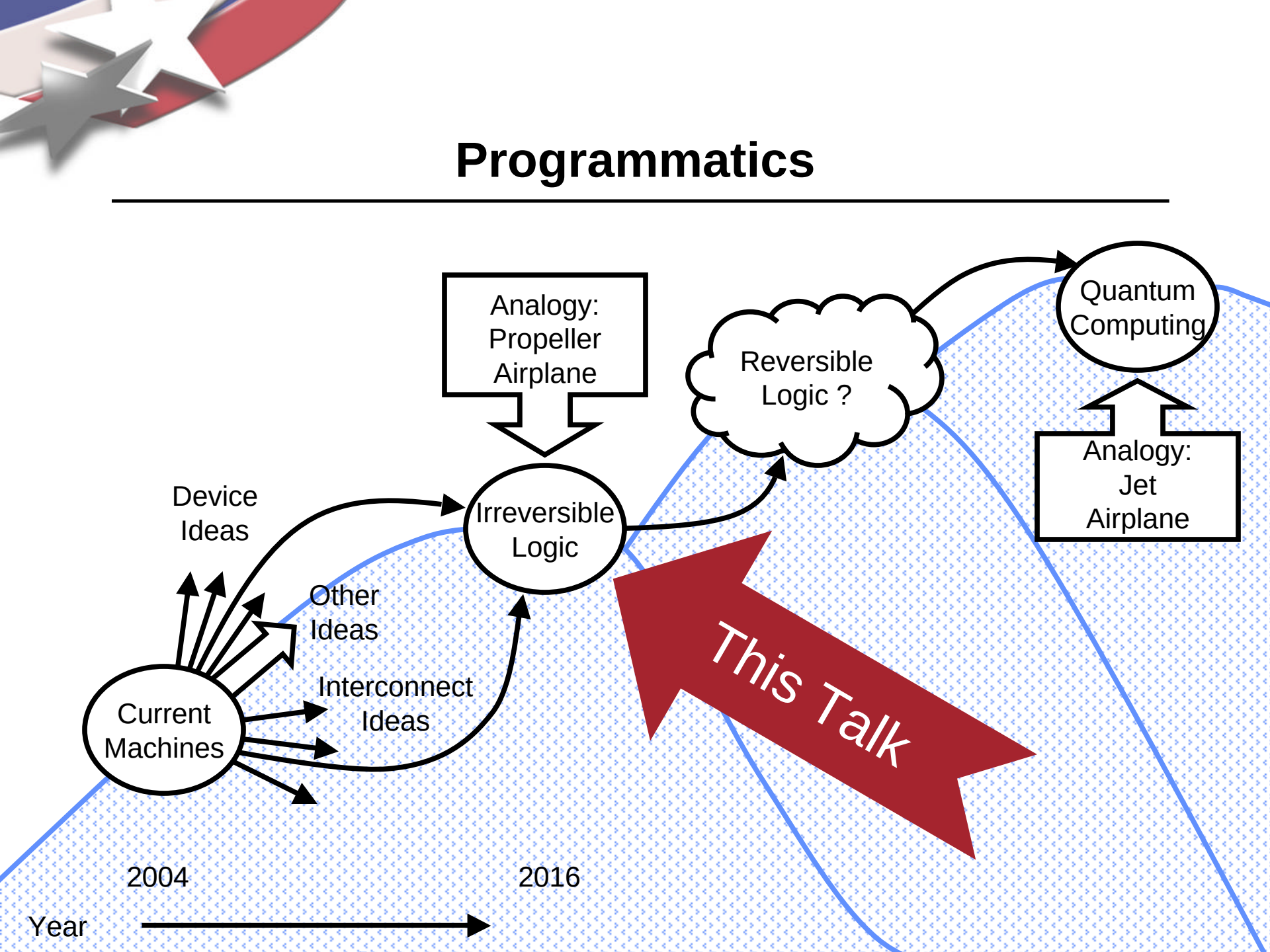


- **We will eventually figure out how to engineer a computer to meet the physical limits**
  - **Semiconductor roadmap shows us at the physical limits about 2016**
  - **Argument applies only to classical irreversible logic with floating point**
  - **Like the engineering of a prop-driven airplane**

Programmatics

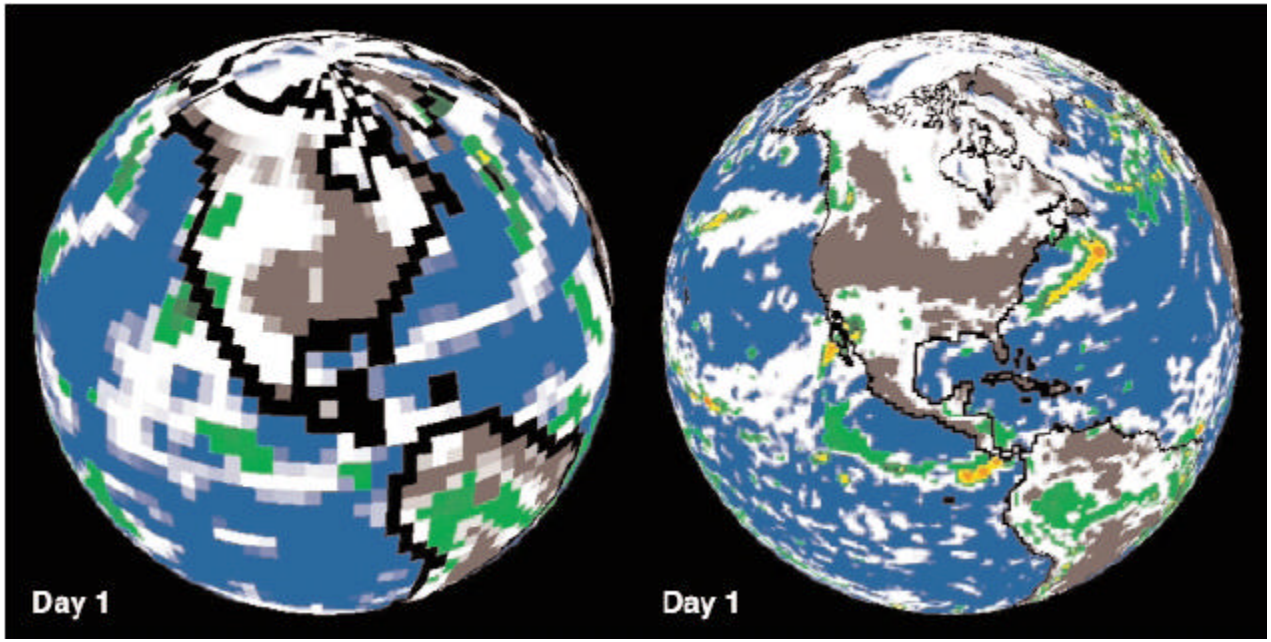


Programmatics



Applications

- **Predominate application is simulation of physics in a computer**



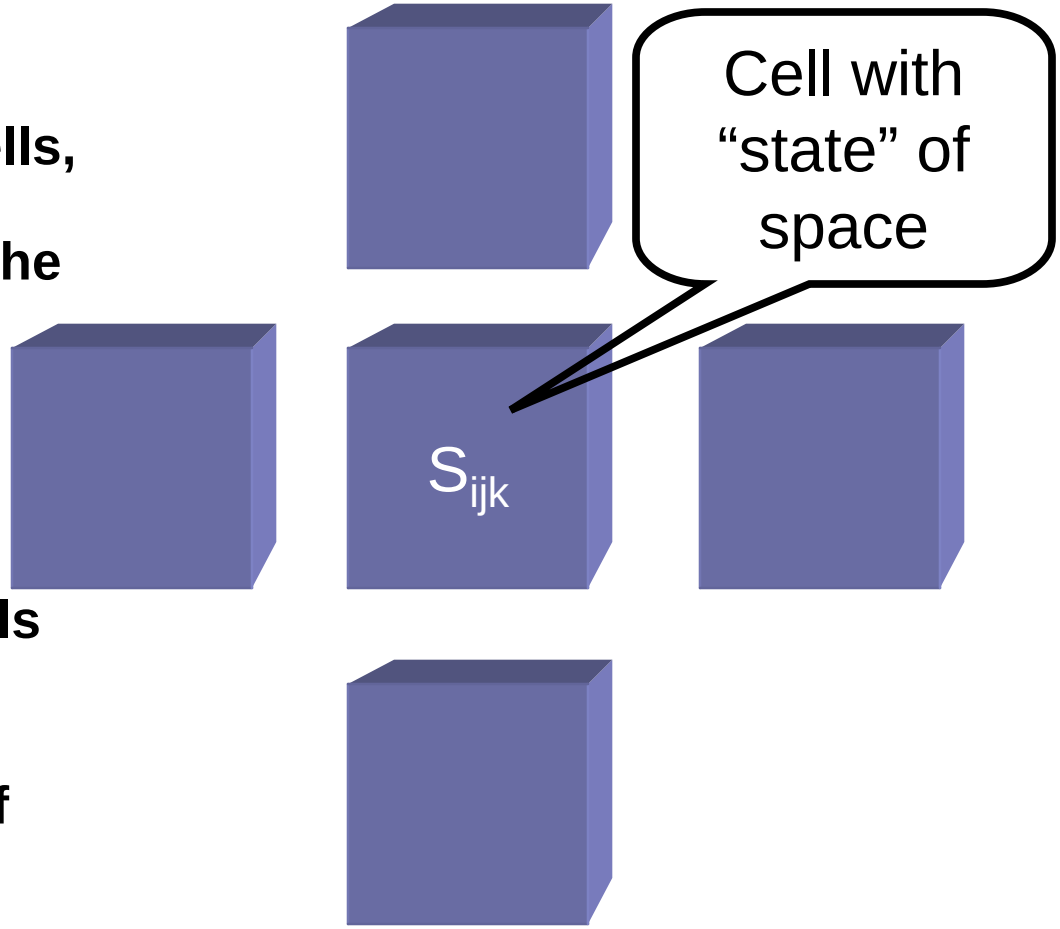
DOE Greenbook

See Richard P. Feynman, “Simulating Physics with Computers,” in Feynman and Computation, A. Hey, ed., Westview Press, 2002.

Sandia National Laboratories

Simulation Algorithm

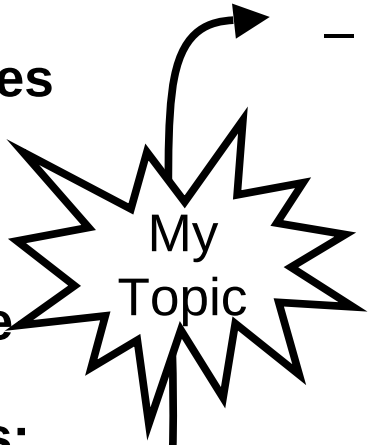
- **Space is divided into cells, each with computer variables representing the physical state of the volume represented by the cell**
- **The computer updates the state of a cell for successive time intervals ΔT based on some physical laws**
- **I. e. S_{ijk}' = f(S_{ijk}, states of nearby cells)**



Sandia National Laboratories

Importance

- When I was in school in the 1980s:
  - Scientific processes consisted of experimental and theoretical
  - We considered the following types of parallel algorithms:
    - Sorting
    - Dense matrices
    - Physical simulation
    - Graphics



- In 2004
  - Physical simulation became a third process in science on equal footing to experiment and theory
    - Dense matrices are a subroutine
  - Graphics and sorting are in parallel commercial computing

1985: 18 Orders of Magnitude Away

- Each irreversible function must dissipate $k_B T \log_e 2$ of heat
- k_B is Boltzmann's constant of 1.38×10^{-23}
- T is temperature in °Kelvin
- $\log_e 2$ is about .7

- Limit for a VAX?

$$\frac{\text{1000 watts}}{k_B T \log_e 2} = 3.5 \times 10^{24}$$

irreversible functions per second

- 10^{18} above VAX's 10^6 operations per second

Irreversibility and Heat Generation in the Computing Process

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

Sandia National Laboratories

Landauer's Arguments

- Landauer makes three arguments in his 1961 paper
  - Kintetics of a bistable well
  - Entropy generation
- We review the second →

- Entropy of a system in statistical mechanics:

$$S = k_B \log_e(W)$$

W is number of states

- Entropy of a mechanical system containing a flip flop in an unknown state:

$$S = k_B \log_e(2W)$$

- After clearing the flip flop:

$$S = k_B \log_e(W)$$

- Difference $k_B \log_e(2)$

Metaphor to FM Radio on Trip to Silver City

- You drive to Silver City listening to FM radio
- Music clear for a while, but noise creeps in and then overtakes music
- Why?
  - Signal at antenna weakens
  - Thermal electron noise constant at $k_B T$
- Analogy: You live out the next dozen years buying PCs every couple years
- Electrical effect
  - Moore's Law causes switching energy of gates to decrease at about 30% per year
  - Thermal electron noise constant at $k_B T$

Our Expectations of Reliability

- What is the consequence of a computer making a spontaneous logic error?
  - We replace the computer
  - Worse than DRAM, where we would add ECC logic
  - Less severe than a heart-lung machine, where we would not build the machine in the first place
- A supercomputer operating at the physical limits a dozen years from now will perform 10^{30}-10^{40} gate operations in its lifetime
- To avoid premature replacement, the probability of a glitch in a gate should be 10^{-30}-10^{-40} per operation

Impact of Power on Reliability

- According to the ITRS roadmap, gates in 2016 and based on 22 nm transistors will be at 10× the power necessary to maintain reliable operation
- However, signal energy is lost for all sorts of reasons and manufacturing tolerances make it unwise to design to the limits
- End of road is on the map!

SNR (db)	Power ratio	P_{error}
10	10	3.90E-06
12	16	9.00E-09
14	25	6.80E-13
16	40	2.30E-19
18	63	1.40E-29
20	100	1.0E-45 – Digital Limit
22	160	3.30E-71
24	250	1.40E-111
26	400	1.80E-175
28	630	1.10E-276
30	1000 – 2016 CMOS	4.5E-437
32	1600	3.5E-691
34	2500	7.1E-1094
36	4000	4.9E-1732
38	6300	2.2E-2743
40	10000	3.2E-4346
42	16000	1.8E-6886
44	25000	1.8E-10912
46	40000	3.8E-17293
48	63000	8.3E-27406
50	100000 – Current CMOS	3.2E-43433

Floating Point

- A floating point unit has about 100,000 gates
- About 20,000 gates will switch for each operation
- Therefore,

$$E_{FLOP} \approx$$
$$20{,}000\ E_{gate} \approx$$
$$2{,}000{,}000\ k_B\ T$$

- Landauer limit is:

  100 TFLOPS/watt

- Accounting for engineering losses, more realistic:

  10 TFLOPS/watt

- If a μP is .1% efficient, the probable limit for a microprocessor is:

  1 TFLOPS/100 watts

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_{sdleak} (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/transconductanc	30%	70%	100%
Parasitic source/drain resistance (Rsd) (ohm-μm) [9]	110	90	80
Parasitic source/drain resistance (Rsd) percent of ideal ch	25%	30%	35%
Parasitic capacitance percent of ideal gate capacitance [1	31%	36%	42%
High-performance NMOS device τ (C_{gate} * V_{dd}/ I_{dd}-NMOS)(ps) [12]	0.3	0.22	0.15
Relative device performance [13]	4.3	7.	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

1000 $k_B T$

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known

Sandia National Laboratories

Global Warming Problem

- Simulate Earth's Weather/Climate
  - 1 foot cell size
    - Earth's surface to 10 km high
  - .01 seconds time interval
  - 100 years simulated time
  - 100 FLOPs per cell update
- Characteristics of solution
  - 3×10^{11} timesteps
  - 2×10^{20} total cells
  - 6×10^{33} FLOPs
  - 2×10^{18} Joules
- Required power source
  - 1.5×10^{19} Joules/year is USA Electric production
  - 50 days of USA electric power production
- On threshold for international project

How Close Is An Opteron To The Limit?

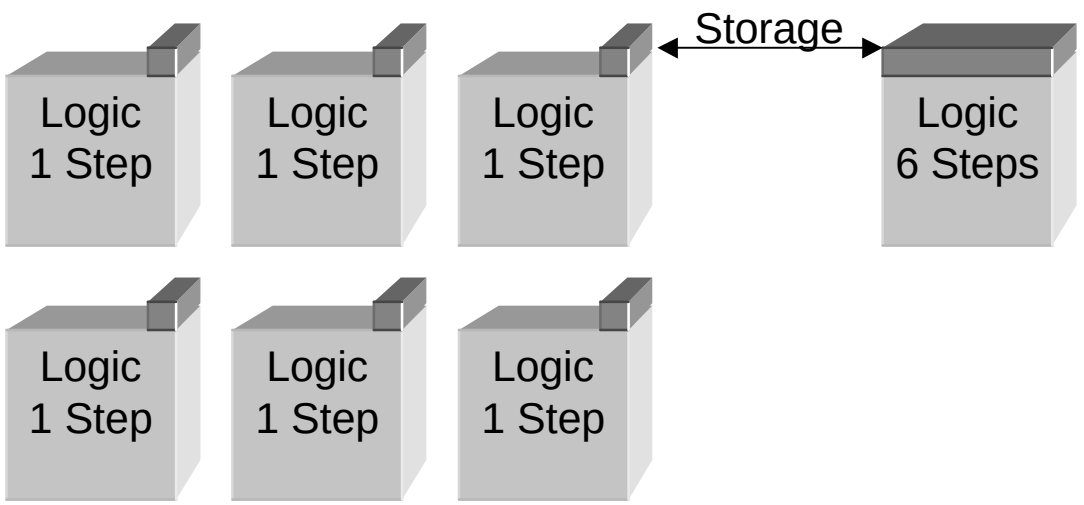
3.48×10^{20}	**Shannon/von Neumann/Landauer limit**
÷ 150	**Noise margin** (100/$\log_e 2$)
÷ 20,000	**Gate operations/FLOP**
÷ 4.5	**Other devices** (.002 fJ → 100k_BT)
÷ 100	**ITRS improvement** (130 nm → 22 nm)
÷ 2	**Opteron voltage** (1.55V instead of 1.1V)
÷ 1500	**Opteron architectural inefficiency**
= 87×10^6	**Opteron performance** 4G/(56W-10W)

Sandia National Laboratories

Time-Space Tradeoff

- **We cannot tell you whether to seek a fast or low-power computer, that is a business decision**
- **But we can explore architectures that trade off speed for low power**

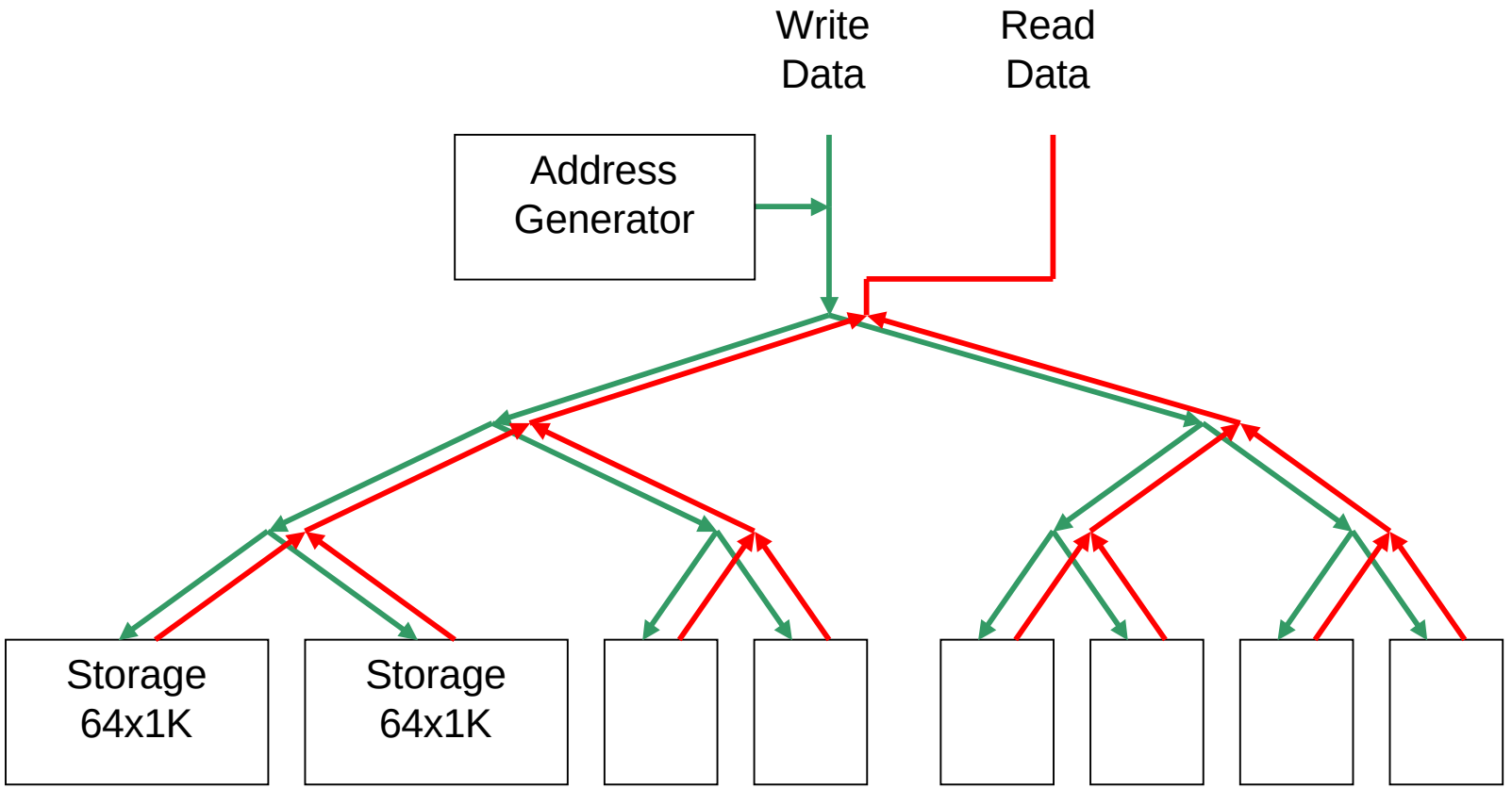
High Speed
High Power
Systolic Array
K=1



Lower Speed
Lower Power
Processor +
Storage
K=6

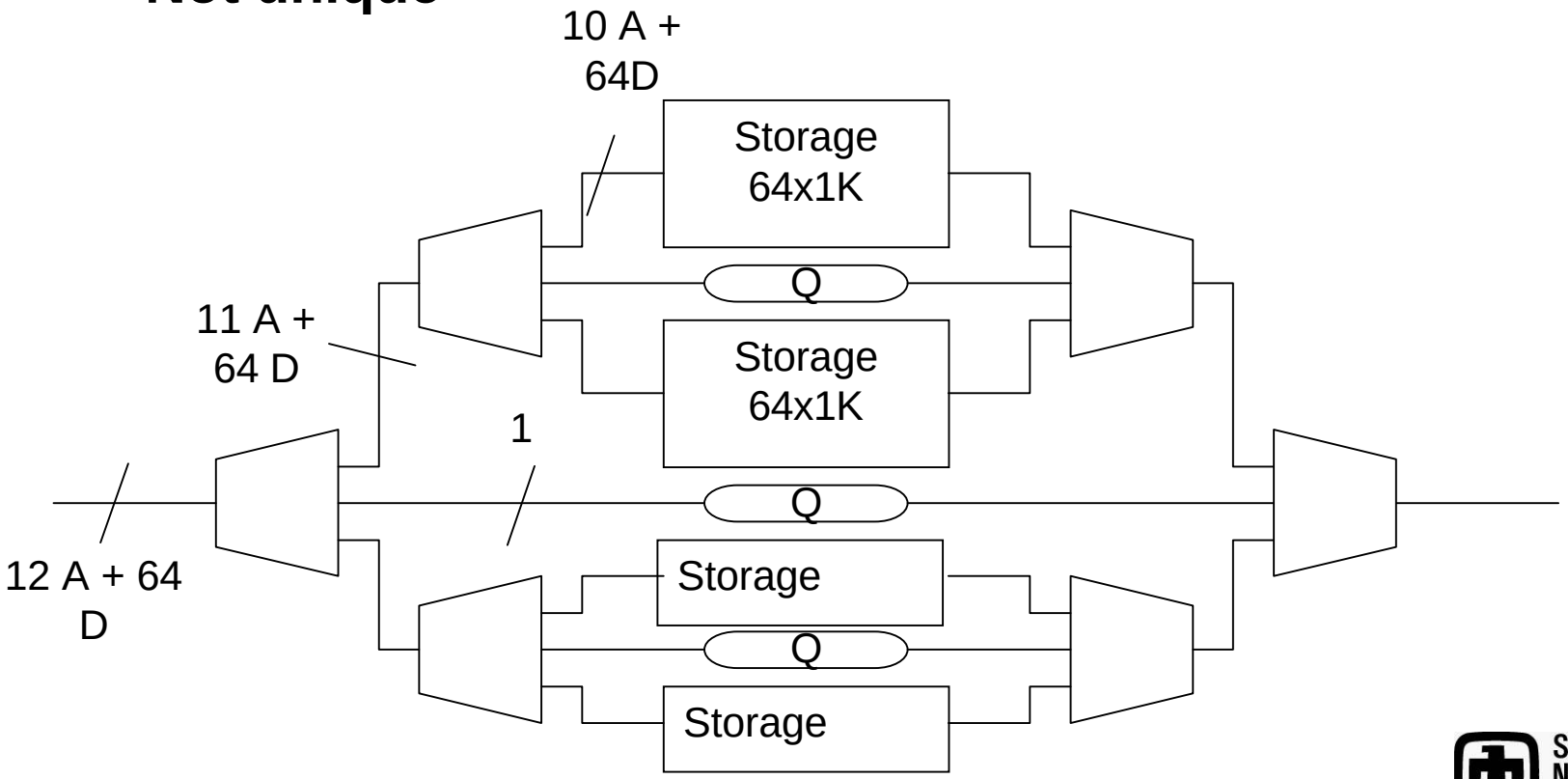
Sandia National Laboratories

O(log size) Memory



Sun FLEETZero

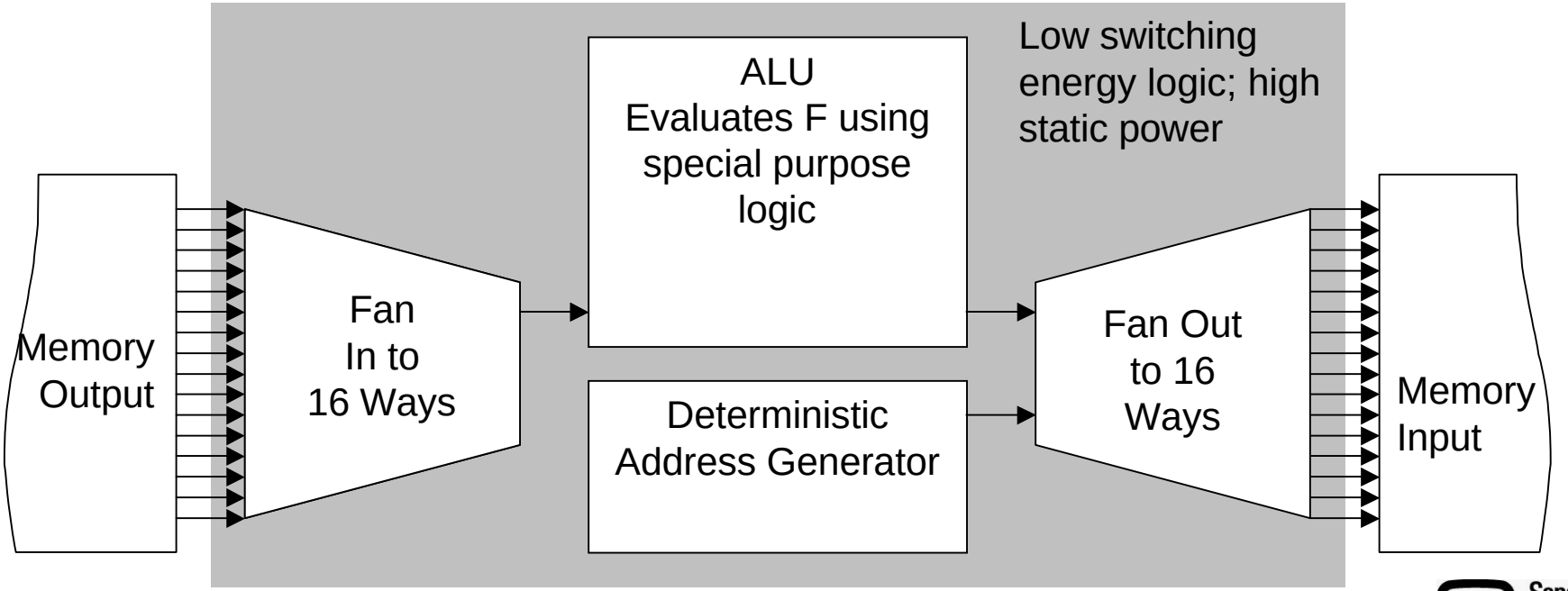
- **Sun FLEETZero has right structure for memory**
  - **Not unique**



Sandia National Laboratories

Special Purpose Logic Solution

- **Address generator generates “raster scan” of cells through space**
- **Memory system must be O(1) power**



Sandia National Laboratories

Application Modeling

- Sample Problem
  - 3D finite difference equation with global synchronization
  - SOR method

$$T_{Step} = \frac{K \times F_{cell}}{floprate} + T_{Global}$$

  - where
    - K is memory size

- Global synchronization limited by speed of light

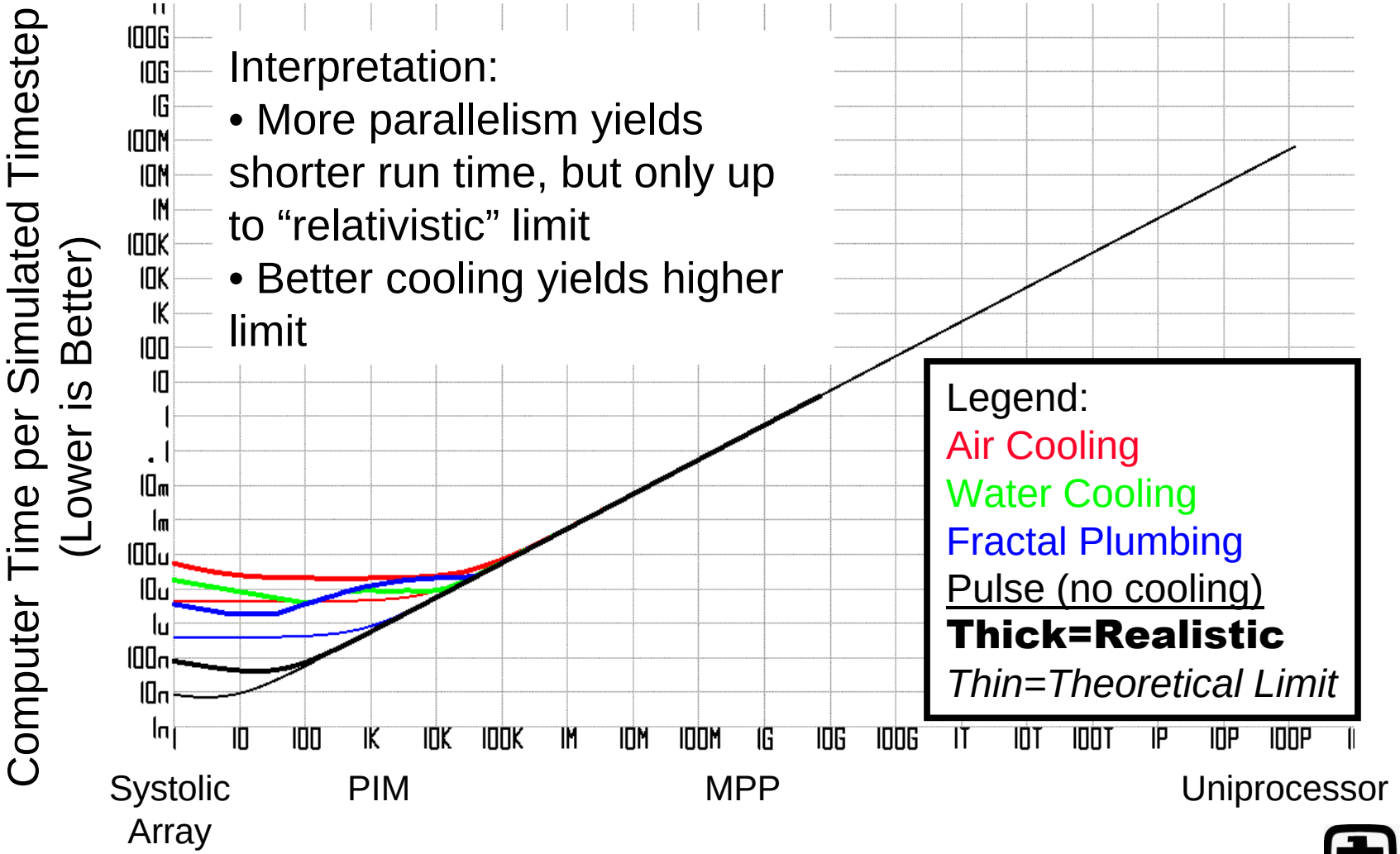
$$T_{Global} \geq \frac{2\sqrt{3} \times L_{Edge}}{c}$$

  - where
    - L_{Edge} is edge dimension of cube

$$6 \times L_{Edge}^{2} \times C_x \leq Power$$

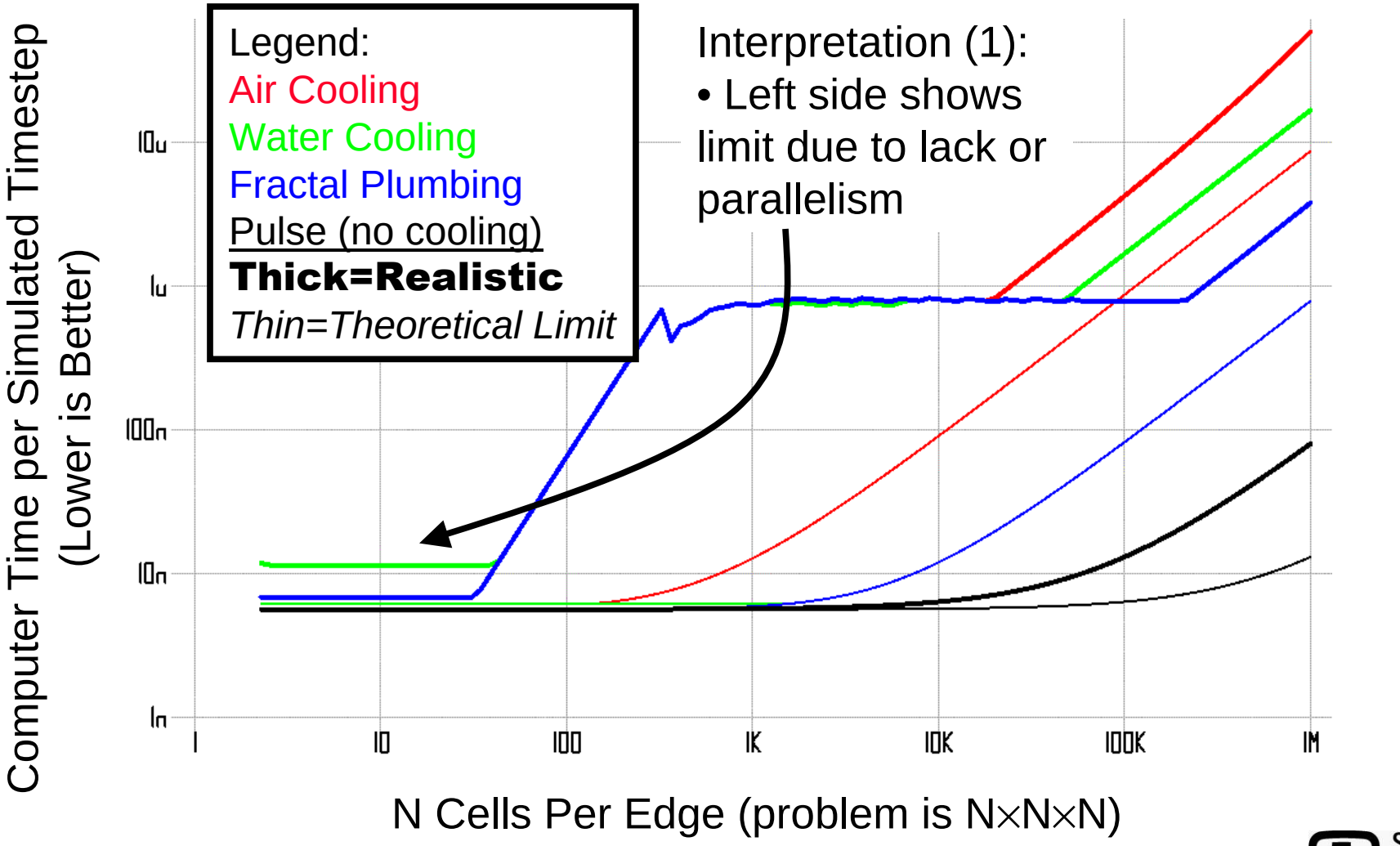
Sandia National Laboratories

Performance on Sample Problem



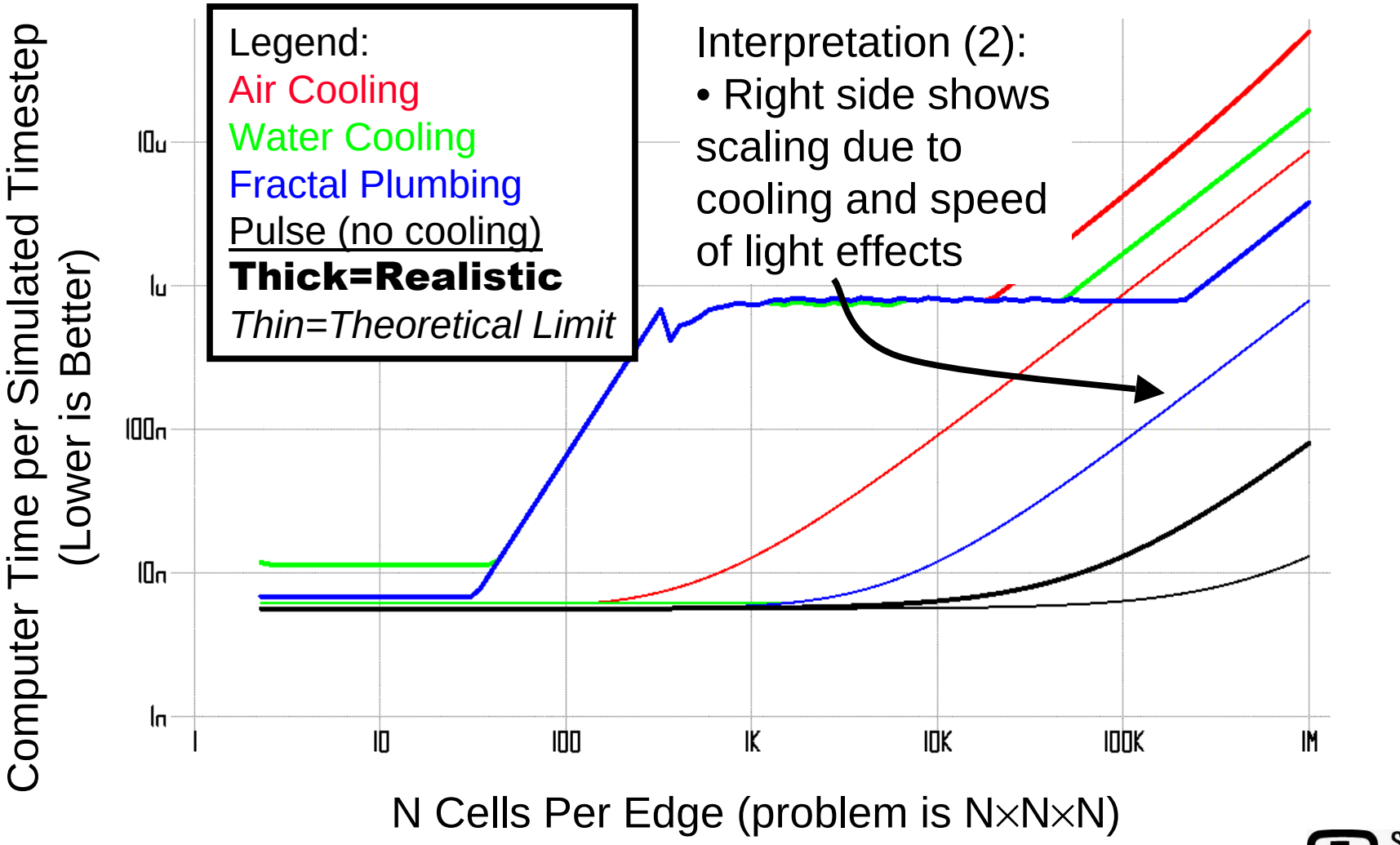
Sandia National Laboratories

Scaling



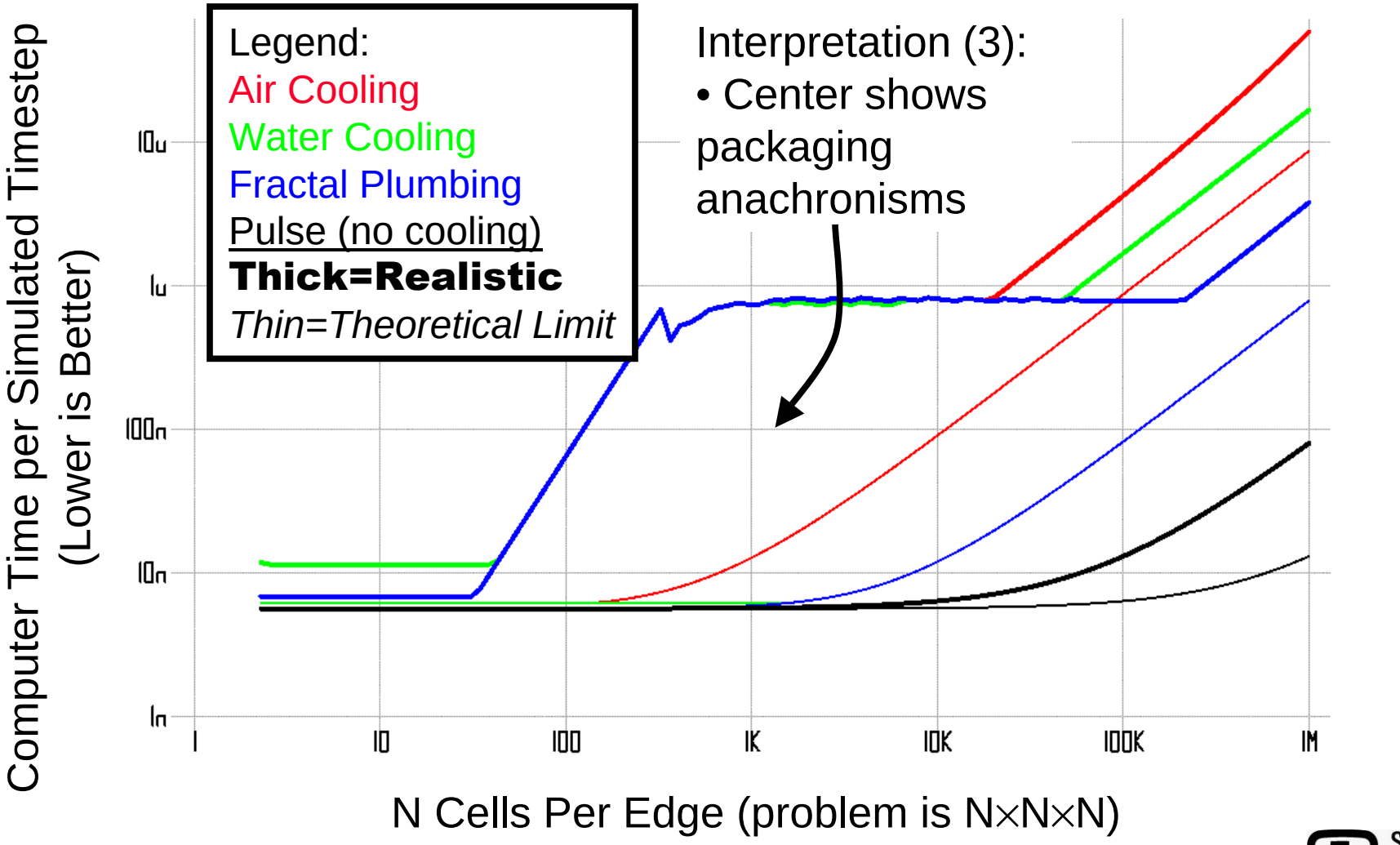
Sandia National Laboratories

Scaling



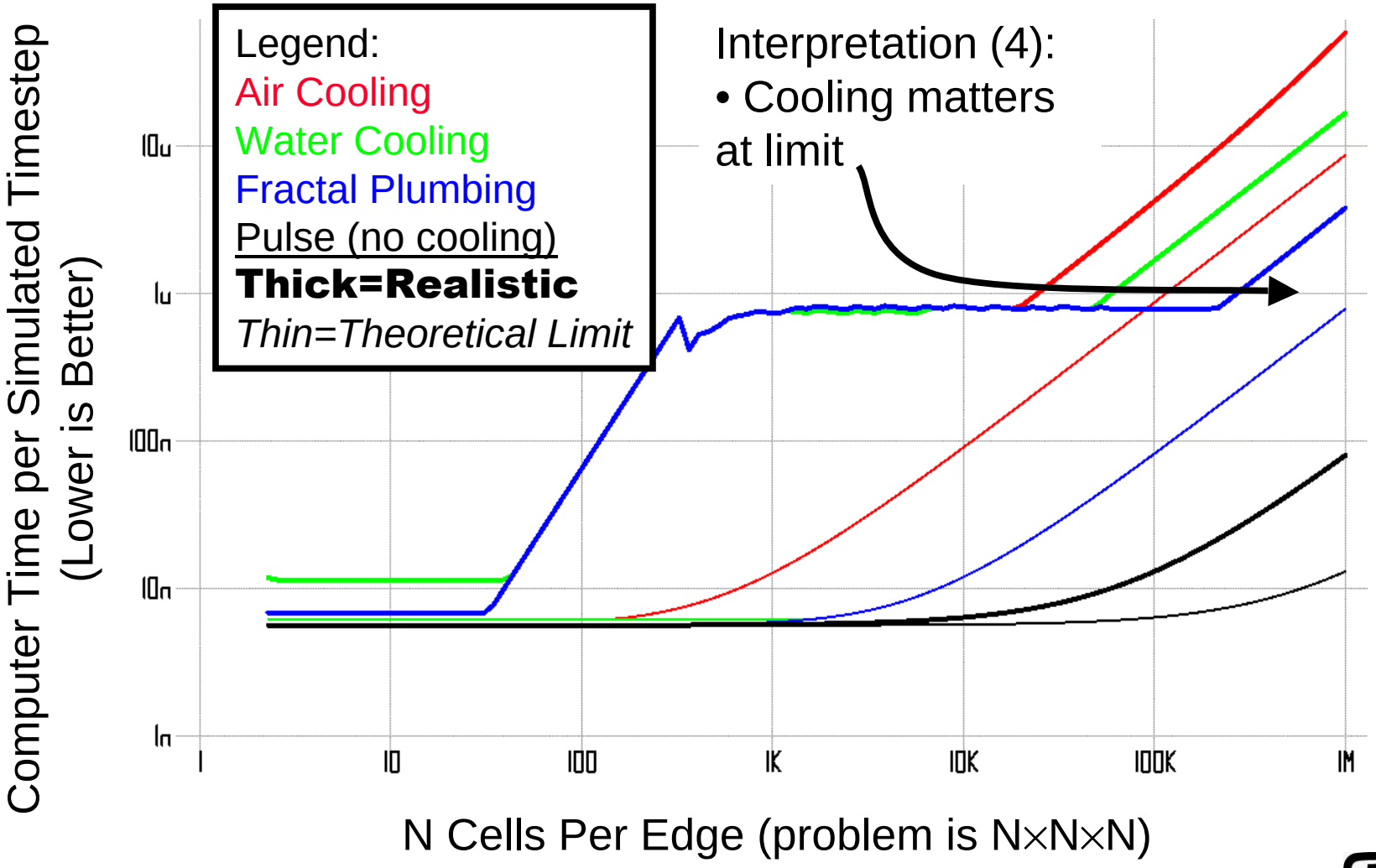
Sandia National Laboratories

Scaling



Sandia National Laboratories

Scaling

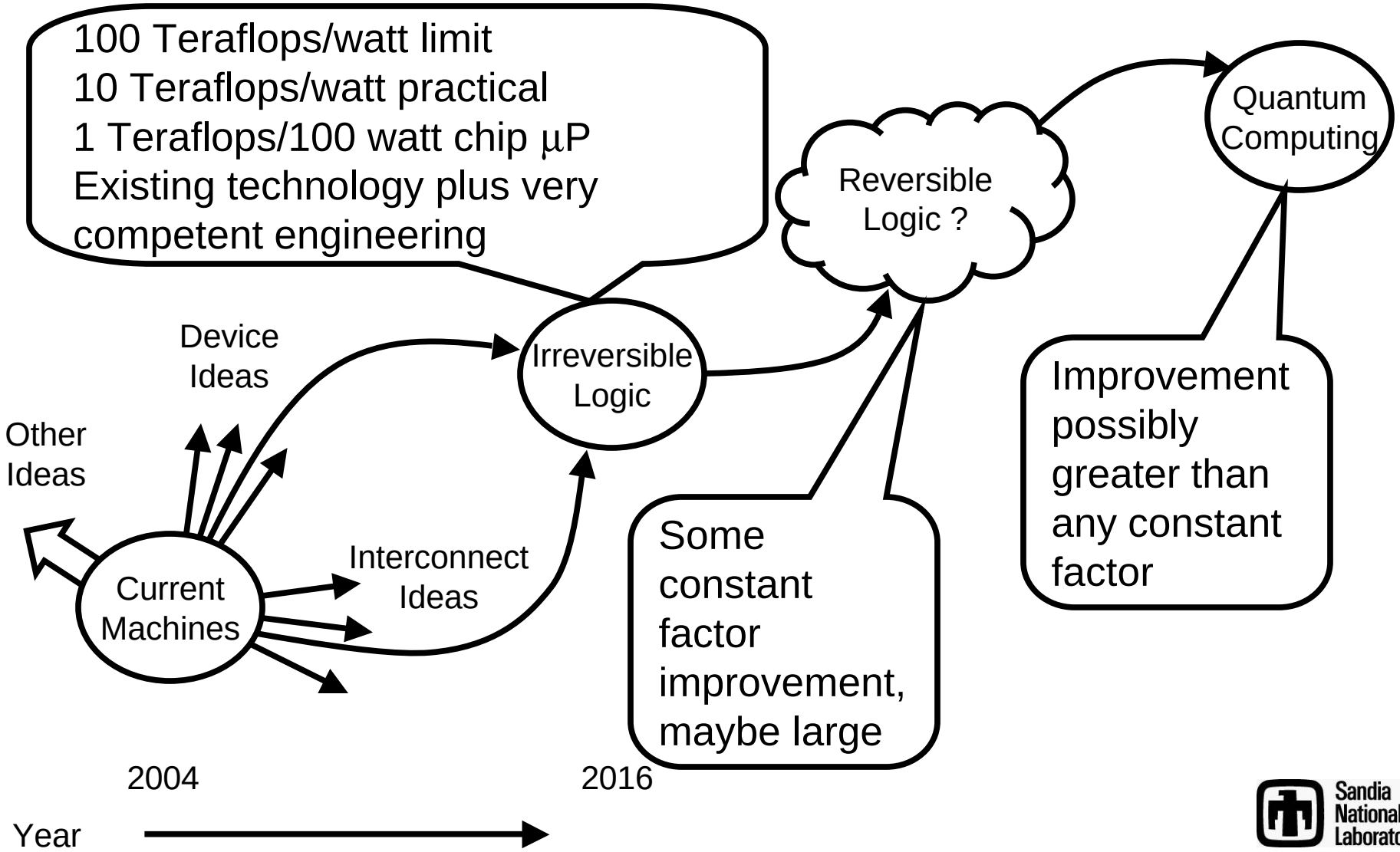


Sandia National Laboratories

General Specifications at Physics Limit

	Red Storm	Limit μP Mode	Limit Turbo Mode
Nodes	10,000	100,000	1,000,000
Node Type	μP	μP	TBD – say 10 vector pipes
Clock	2 GHz	20 GHz	20 GHz
Flops/node	4 GFLOPS	40 GFLOPS	400 GFLOPS
System Peak	40 TFLOPS	4 PFLOPS	400 PFLOPS
MPI Latency	2.5 μS	100 ns	N/A – no MPI
Power	2 MW	2 MW	2 MW

Conclusions



Sandia National Laboratories

Ways To Proceed

- Need architectures that do more per…
  - Watt?
  - Gate?
  - Gate switching operation?
  - Overall, we need to maximize the fraction of the logic gates that are contributing to the useful result of the calculation
- What are the application programs?
  - Science needs large simulations
  - Intelligence community needs code cracking
  - Commercial industry doesn't have really big problems
  - Maybe new things
- What applications require quantum computing, etc.