SAND2006-3728C

# Sandia's Red Storm Supercomputer and the Transformation of Engineering

## June 12, 2006

## New Mexico Society of Professional Engineers

**Rob Leland**

**Director, Computing and Network Services**

**Sandia National Laboratories**

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. SAND -2005-6648P



Sandia National Laboratories

# My objective

- To leave you with answers to three basic questions:
  - What is Red Storm?
  - How is Red Storm being used?
  - Why is Red Storm particularly good at what it does?

# Engineering Transformation at Sandia

- **National drivers**
  1. **Loss of nuclear testing capability**
  2. **Emerging security challenges post 9/11**
  3. **Understanding high consequence events**
  4. **Sustaining economic competitiveness**

# Today: science based stockpile stewardship

- **Central problem**
  - Replacement of underground testing with a rigorous scientific methodology to maintain confidence in our nuclear stockpile.
- **Necessary standard**
  - We must certify annually to the Secretaries of Energy and Defense that the stockpile is safe, reliable and secure.
- **Time urgency**
  - The stockpile is aging, departing further every year from our test experience base.



Sandia National Laboratories

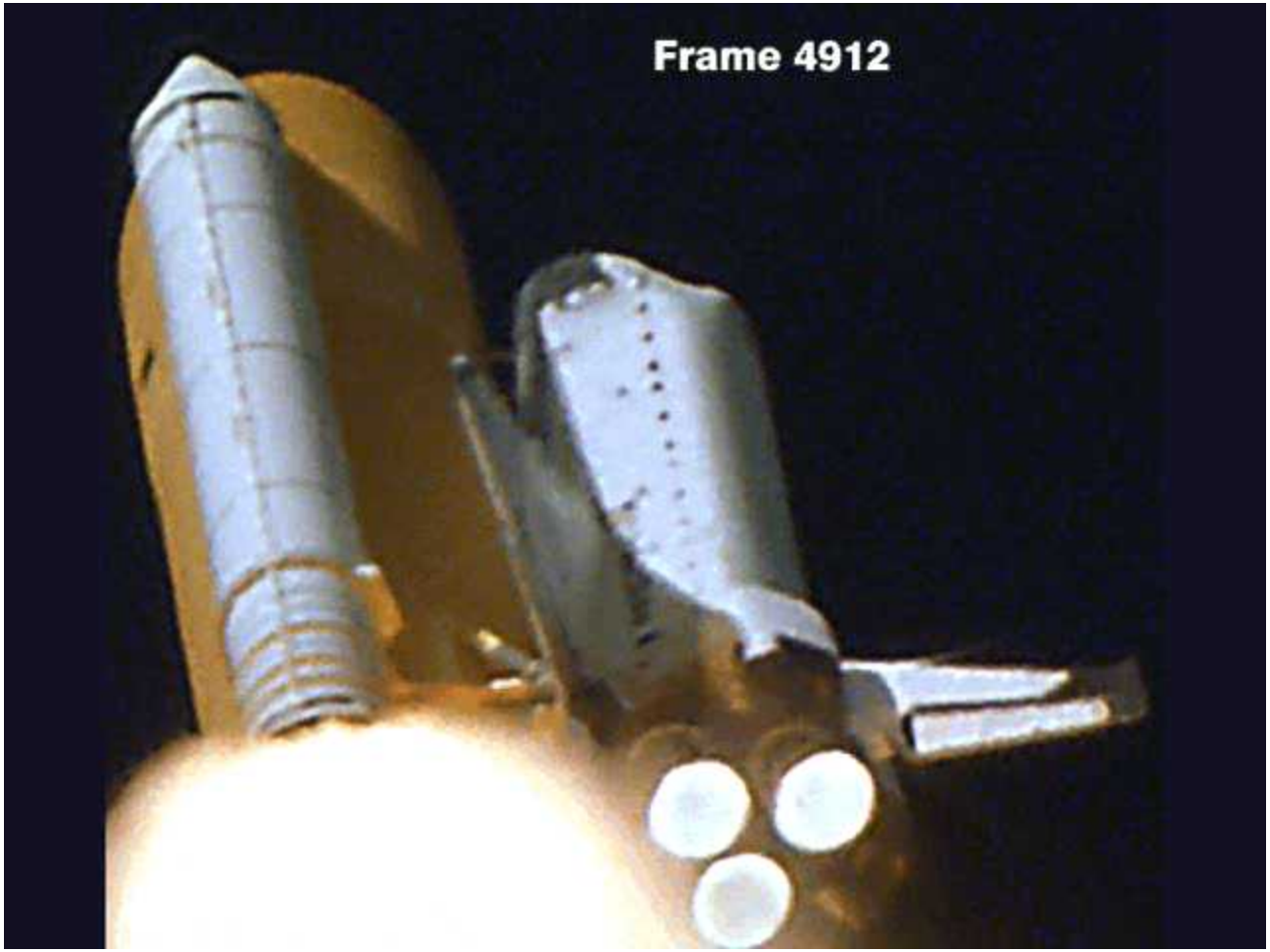
# 2. Emerging security challenges post 9/11

**"… after operations of three years or perhaps less, the Sandia Pulsed Reactor will no longer be needed, since computer simulations will be able to assume its mission."**

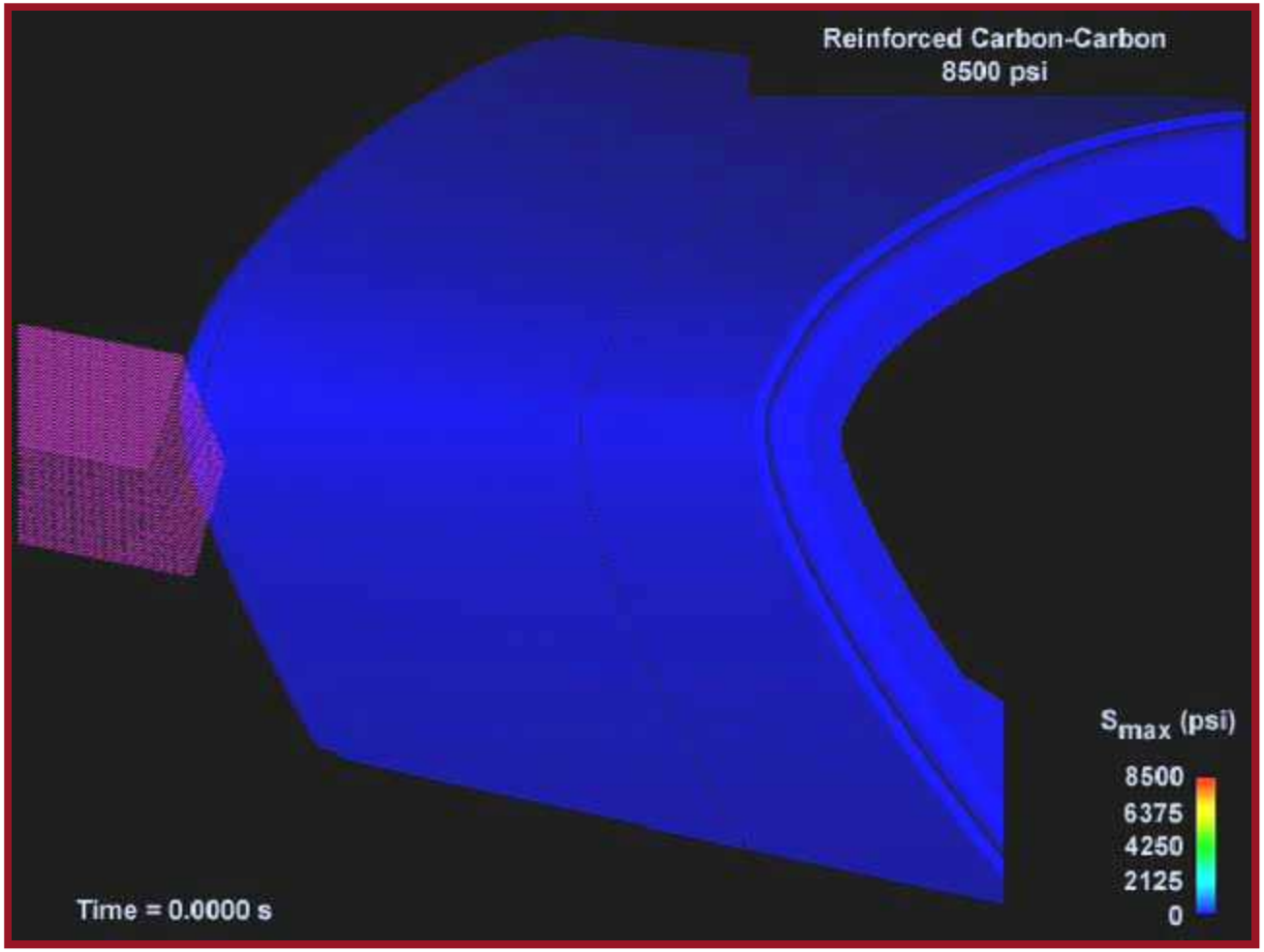
**- Secretary of Energy Abraham, 2004**

- **SPR historically critical to weapon certification**
- **Provides unique combination of high fluence, short pulse width, mixed radiation, and high volume**
- **Being removed from service due to security concerns post 9/11**
- **Replacing SPR implies petascale computation (~25X current capability)**

# 3. Understanding high consequence events



# Simulation of foam hitting shuttle wing



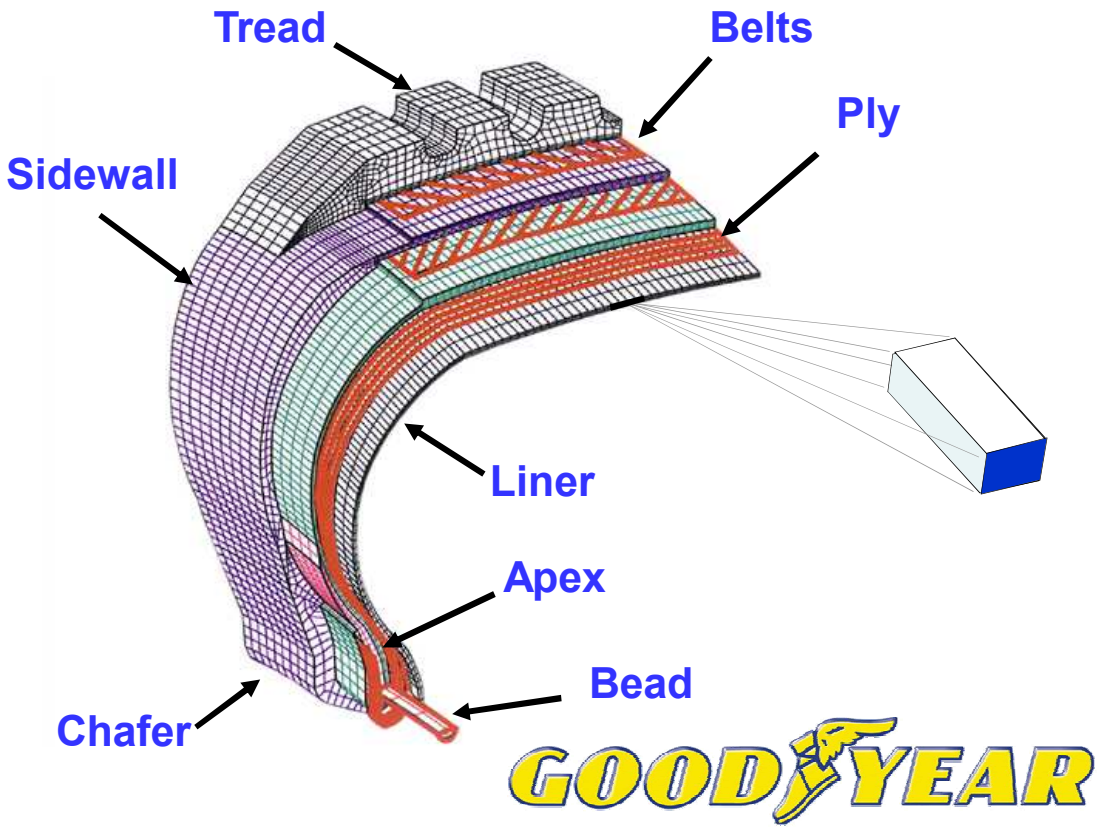



Sandia National Laboratories

# Physical validation test

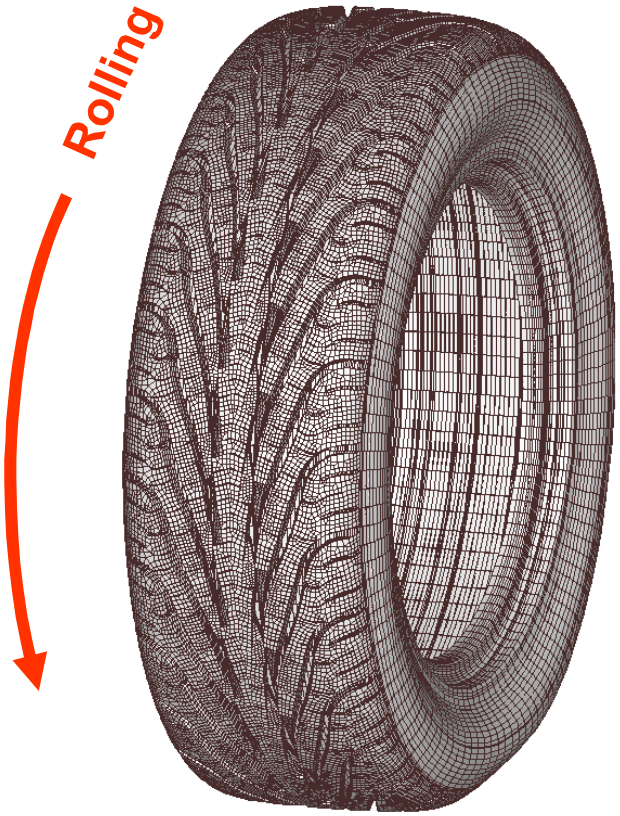
# 4. Sustaining economic competitiveness

## Finite Element Analysis (FEA) Computer Modeling of Tires

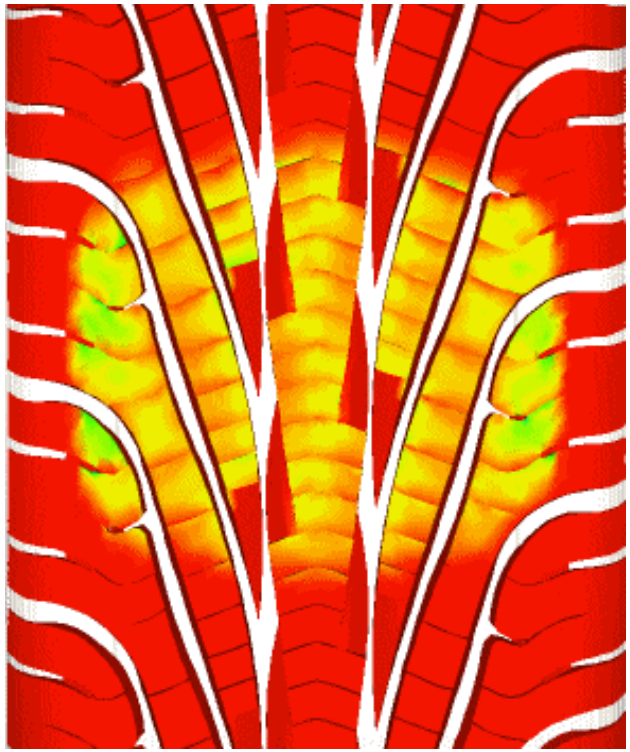



- A tire is a composite structure of at least 8 layers of different materials
- Each layer can be approximated with a large number of "finite" elements
- Most of the elements have non-linear material properties

# Simulation based product design




GOODYEAR



**Wear simulation using Sandia derived tools**



Sandia National Laboratories

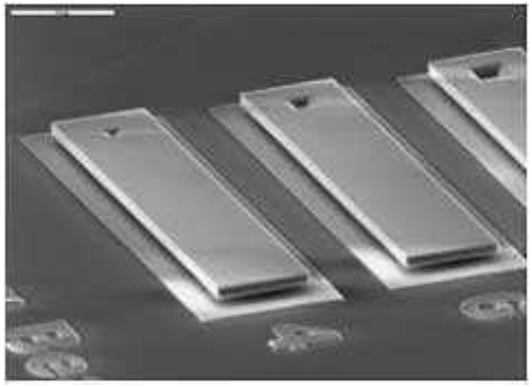
# How much deeper can we go?



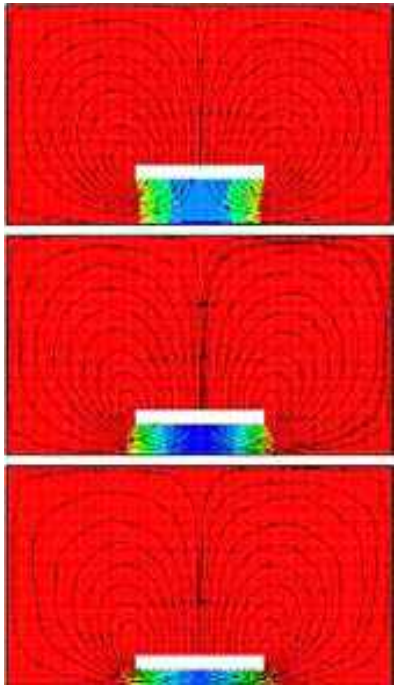

Sandia National Laboratories

# A very “simple” MEMS example

- **Polycrystalline silicon beams**
  - Used in accelerometers, switches etc.
  - Specs: 100 by 20 by 2 microns; 10-100 kHz
  - “diving boards for bacteria”
  - Gas dynamics critical to performance
  - Material properties not homogenous
  - Physical intuition often wrong

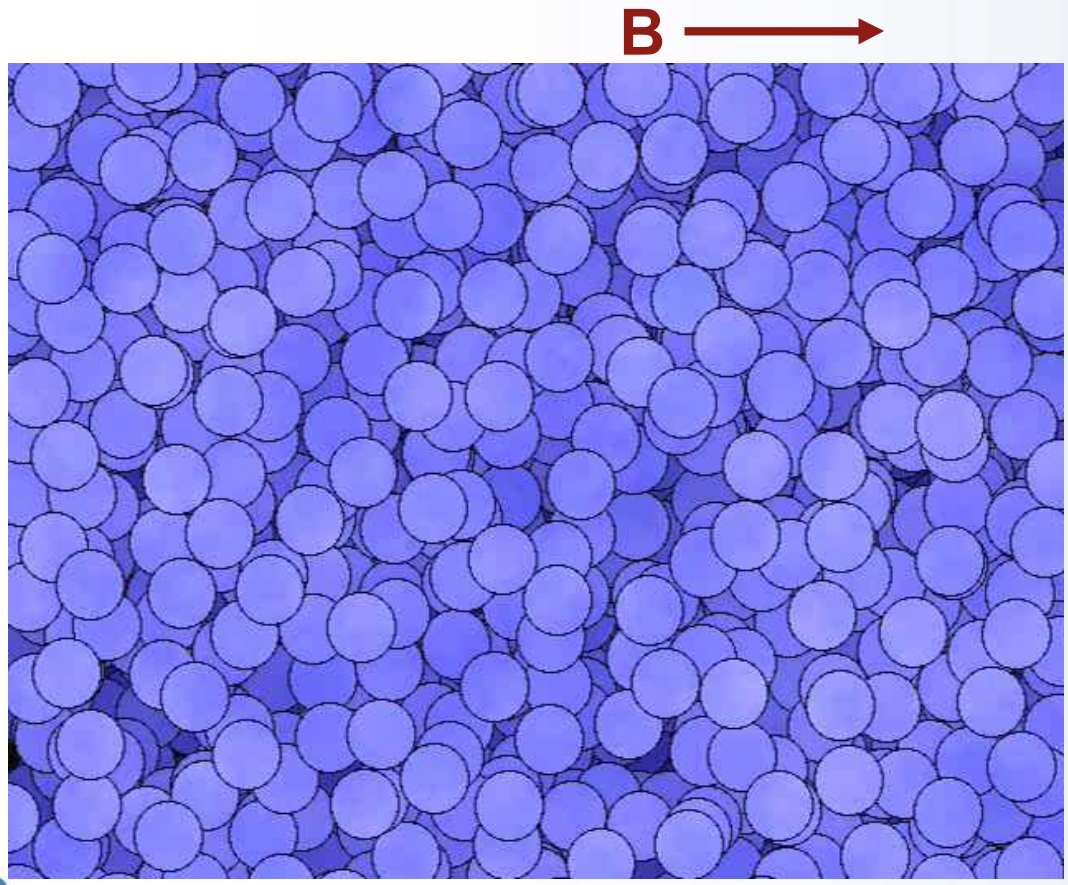


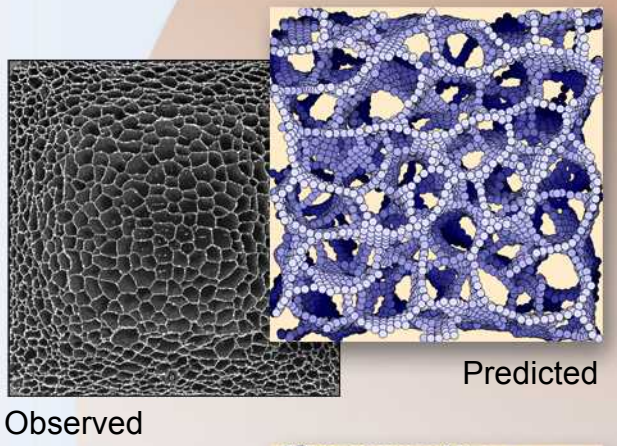
**Problem**
- 3-d geometry
- Moving boundary
- Transient gas flow
- Full oscillation

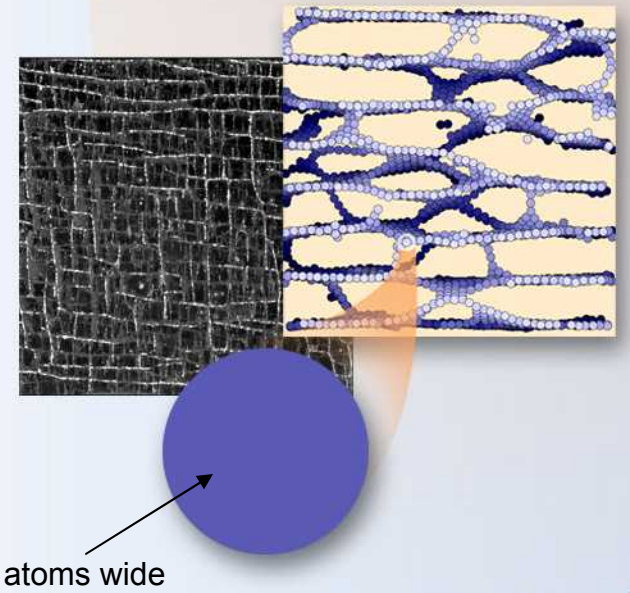


**Current simulation**
- 2-d geometry
- Static boundary
- Steady-state gas flow
- Distinct points in oscillation

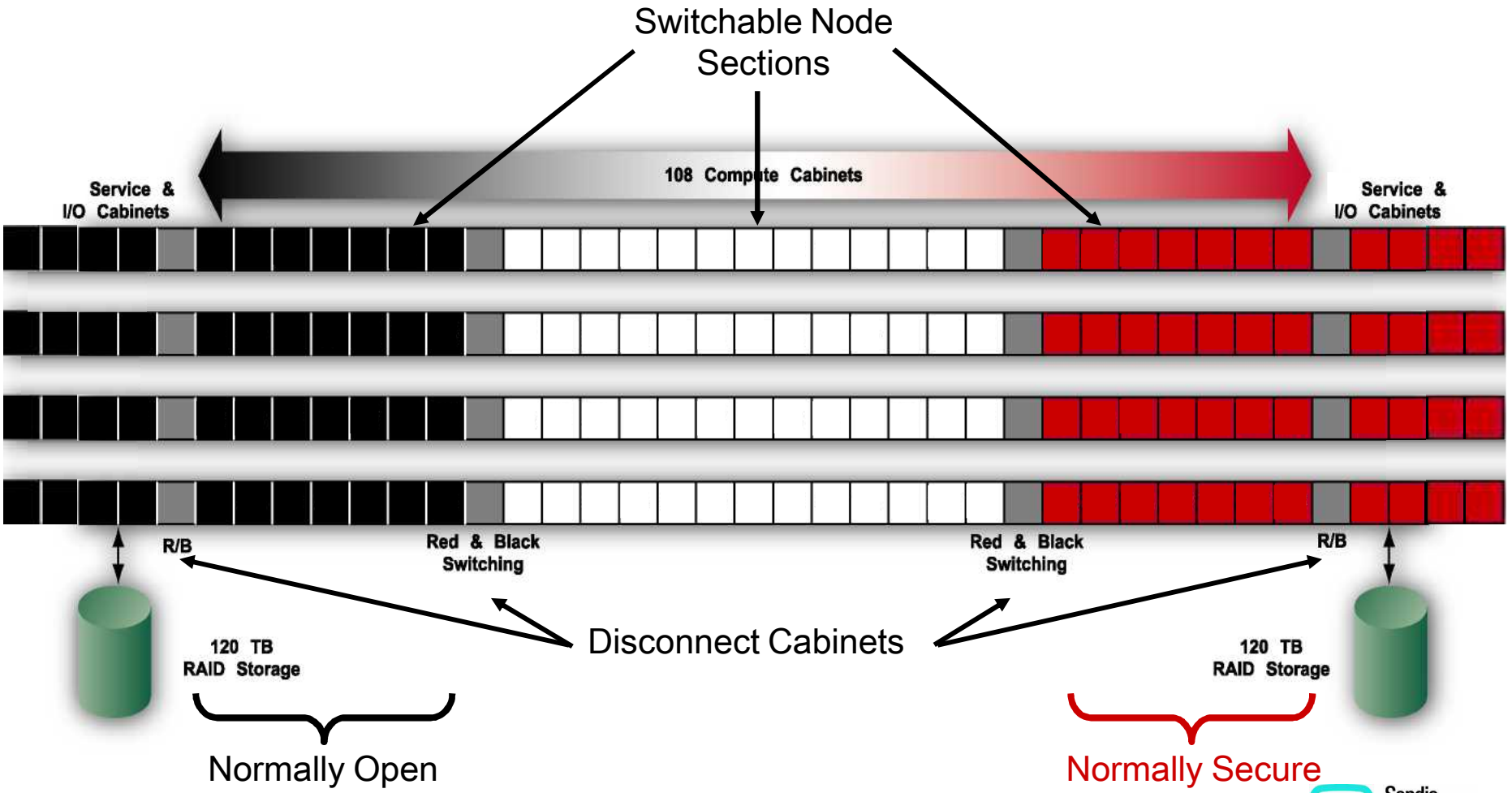
# We can now engineer at the nanoscale



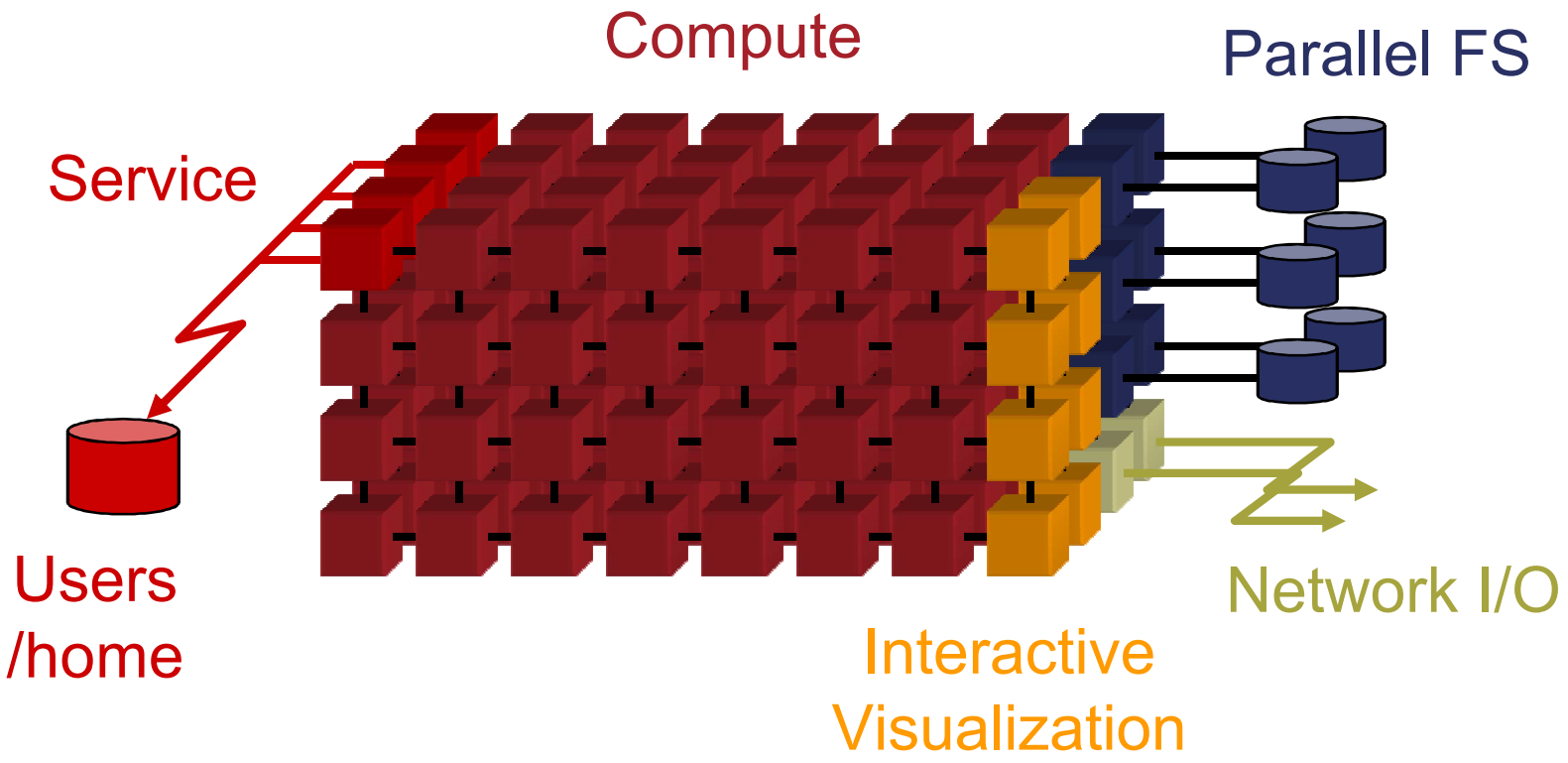
# Red Storm floor plan



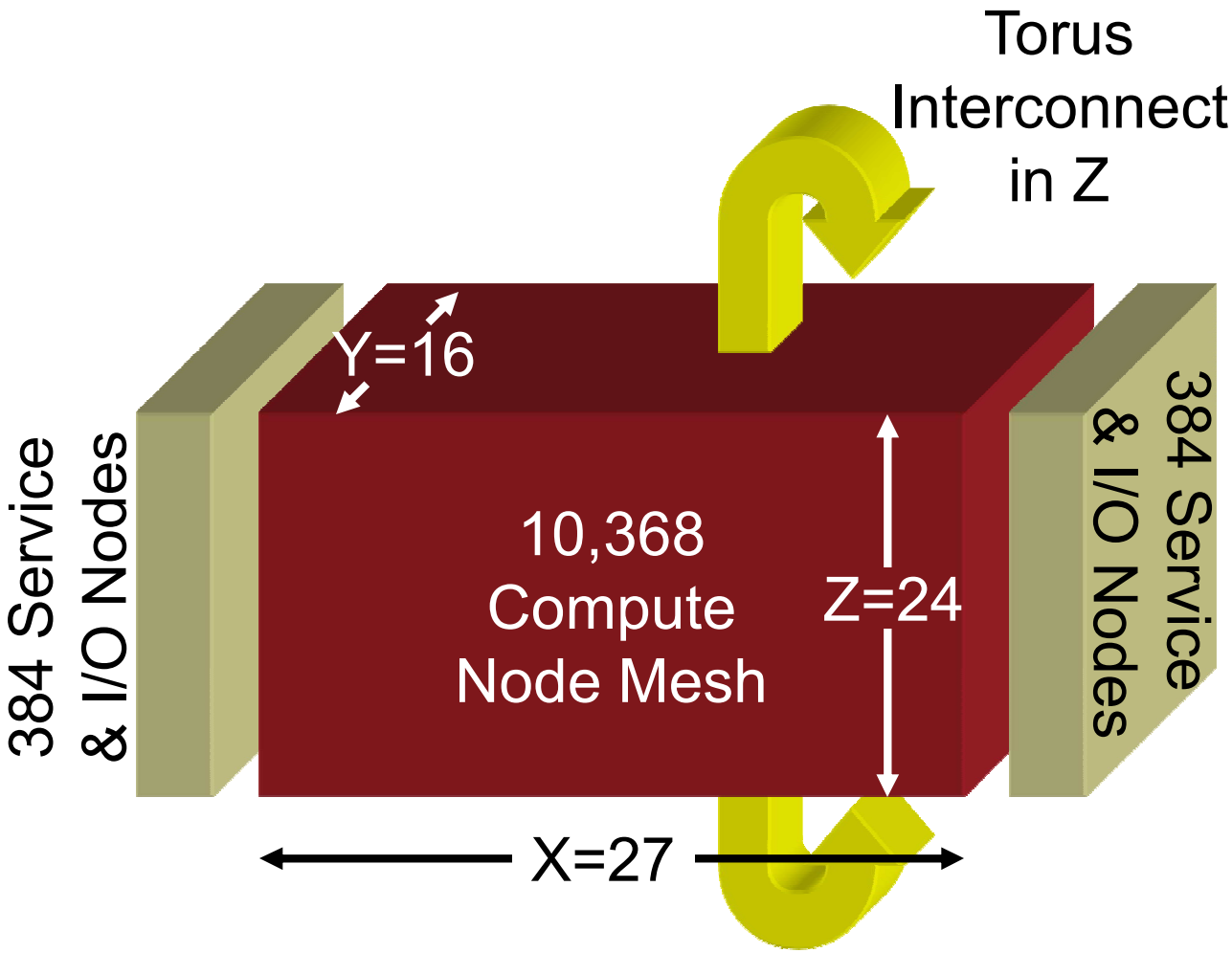



Sandia National Laboratories

# Functional architecture



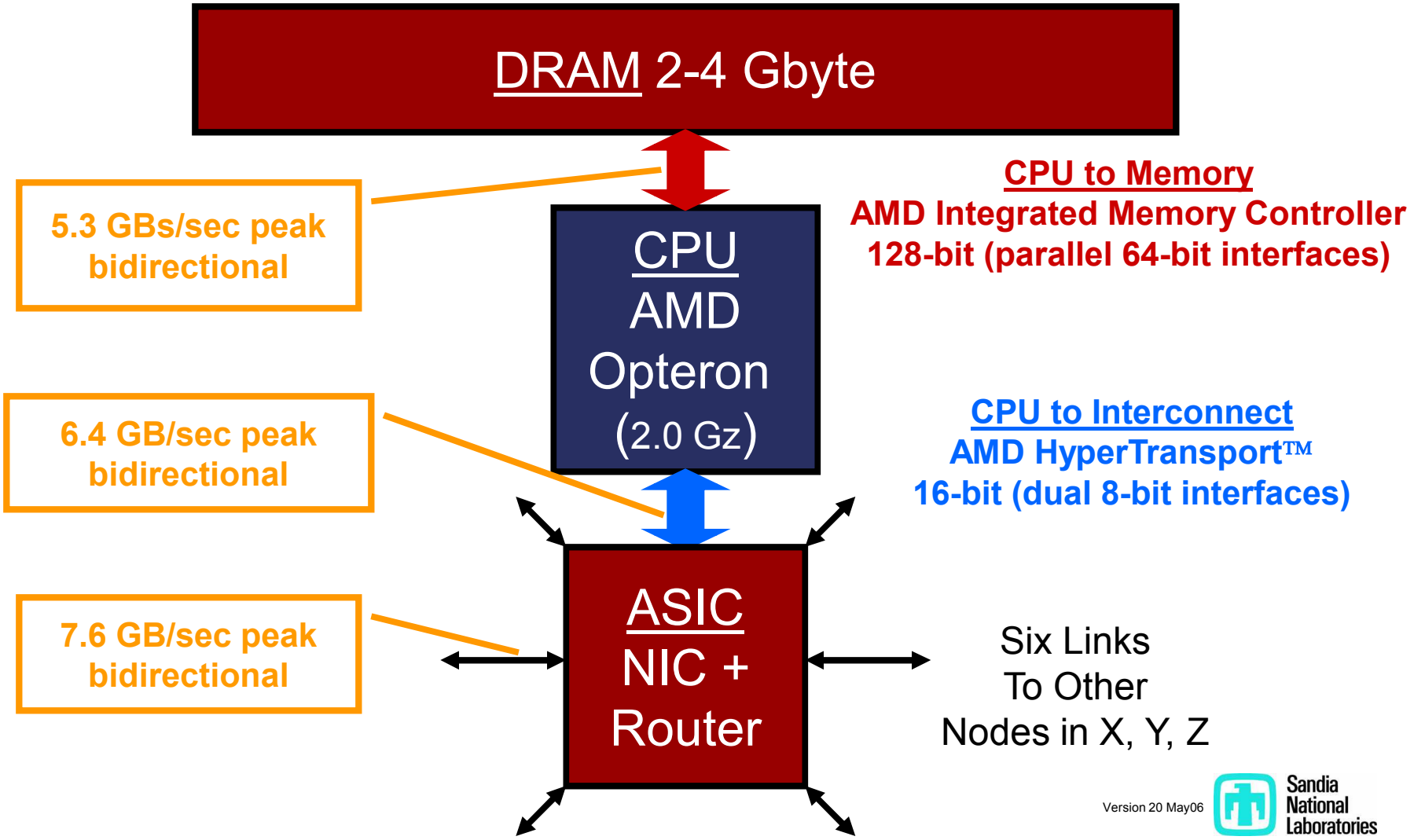
# Topological architecture






Sandia National Laboratories

# Compute node architecture






Sandia National Laboratories

# Key issues in modern supercomputing

- Scalability
- Usability
- Reliability
- Economy

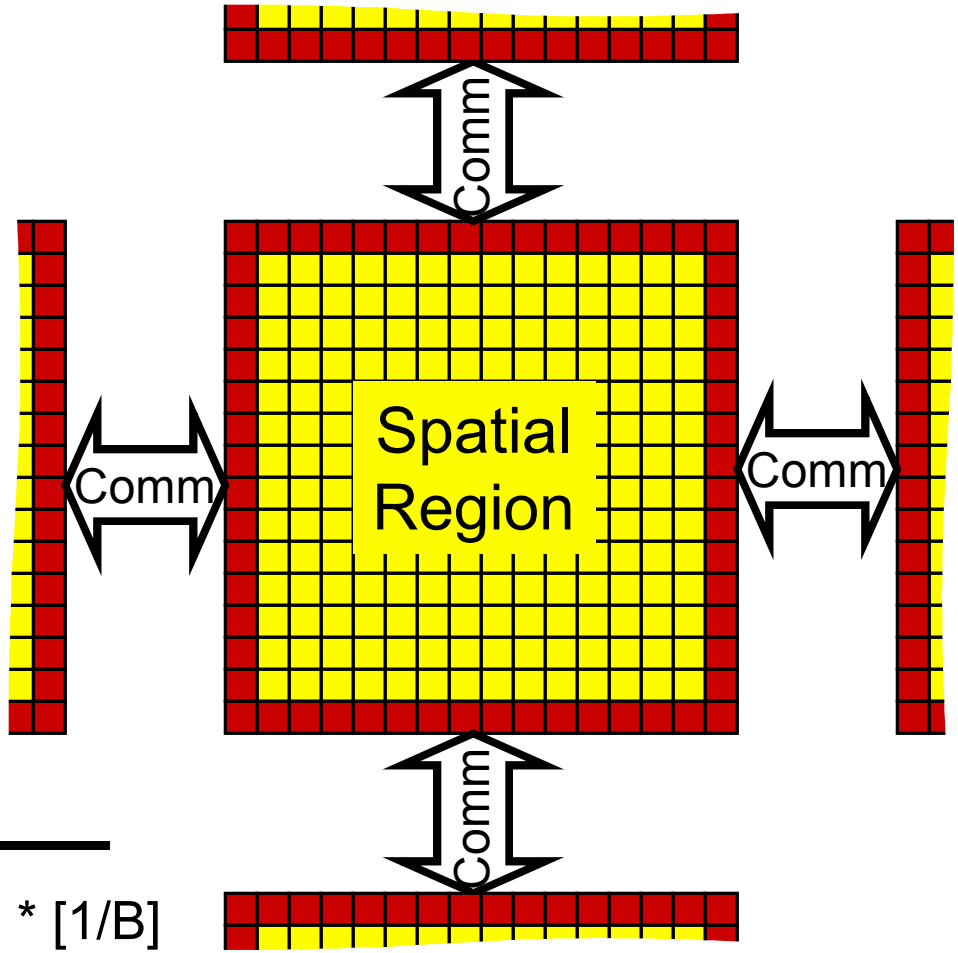
} “SURE” architecture

# Scalability

- Low latency, high bandwidth interconnect
  - Seastar network chip is world's fastest
- Low interference from system software
  - Catamount is Sandia's 4th generation LWK
- Extensible and upgradeable hardware design
  - Mesh topology
  - Thermal & mechanical design headroom
  - Board design
  - Balance "to spare"

# Scalability: a simple model

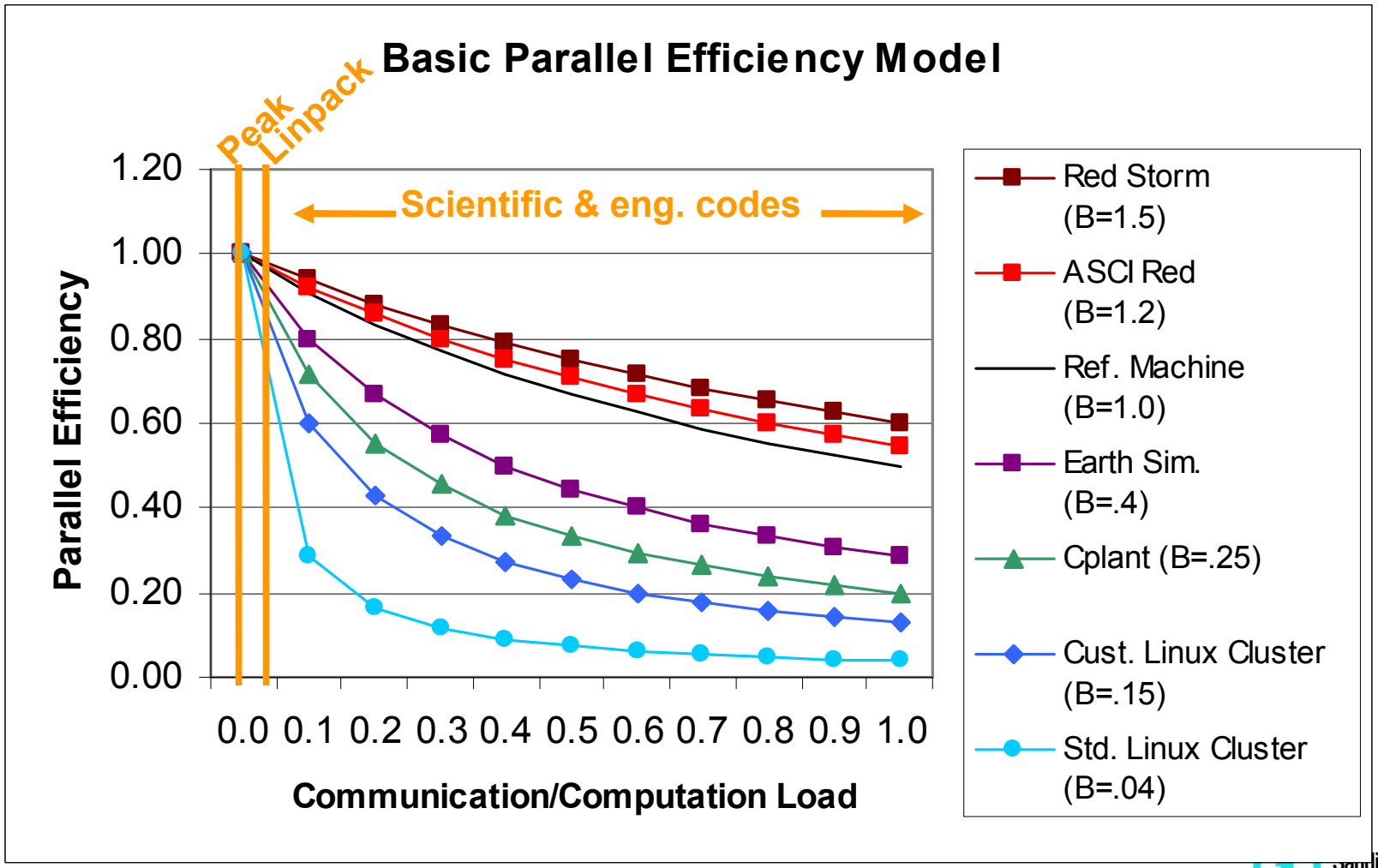
- E.g. nearest neighbor 2D relaxation
- Interior values updated by computation each iteration
- Boundary values must be communicated each iteration
- Overhead is proportional to ratio of boundary to interior points
- Overhead is inversely proportional to communication/computation rate
- Hence efficiency goes like …




$$\frac{1}{1 + \text{overhead}} = \frac{1}{1 + [1/\text{sqrt}(N)] * [1/B]}$$

Sandia National Laboratories

# System balance determines scalability

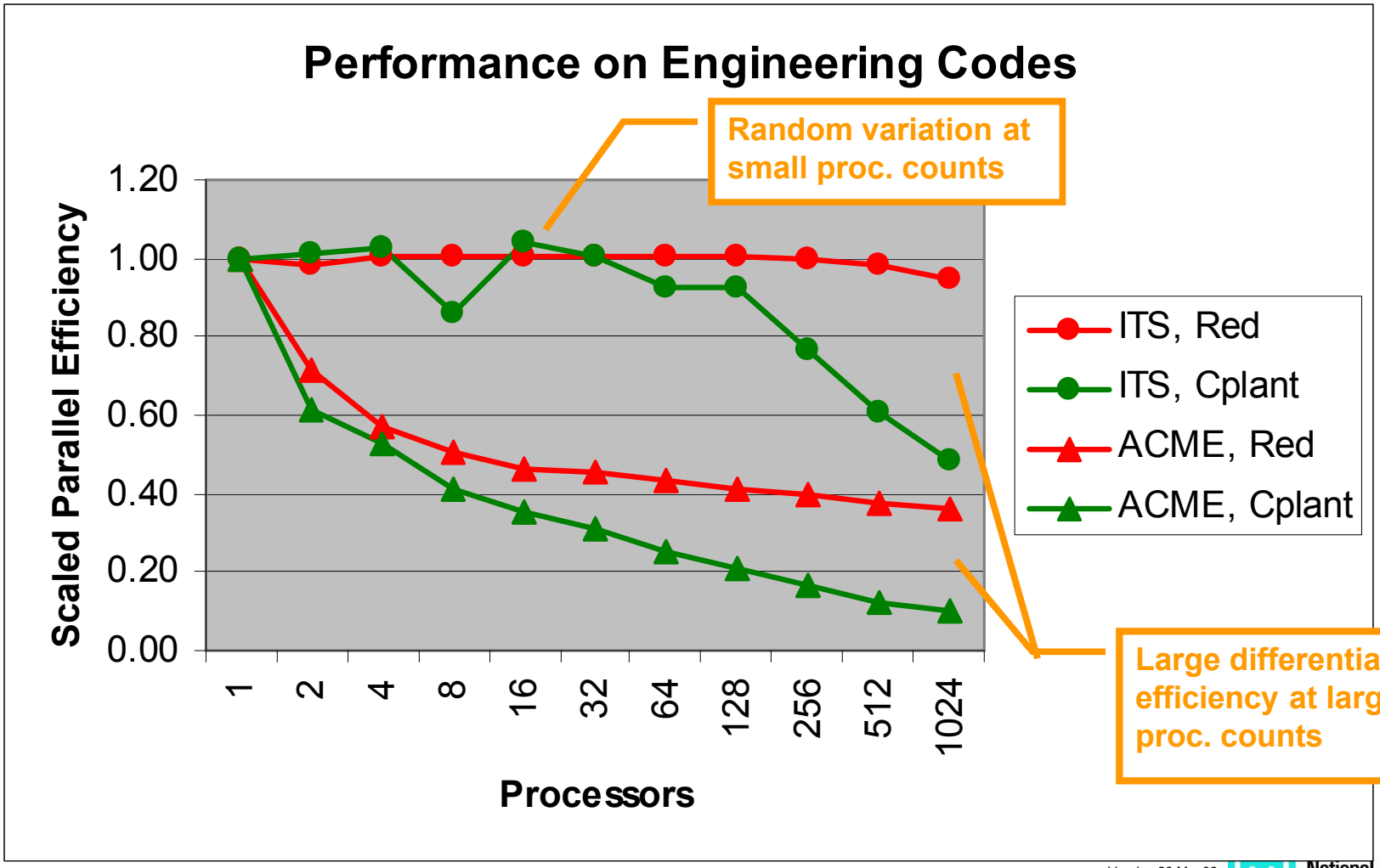


# HPCC benchmarks

| System | HPL | PTRANS | RANDOM | FFT | STREAM |
|---|---|---|---|---|---|
| Red Storm (SNL) | 41 | 100 | 100 | 51 | 79 |
| Blue Gene Light (LLNL | 100 | 19 | 6.5 | 100 | 97 |
| Purple (LLNL) | 72 | 31 | 17 | 39 | 100 |

… percentage of best achieved so far on baseline benchmarks

# Scalability data for some key engineering codes






National Laboratories

# Usability

- Linux front end
- Standard programming paradigm
- Scalable load, boot, allocation, monitoring and IO hardware, algorithms and software
- System reliability …

# Reliability

- NOT high availability
- 100 hrs. MTBI for hardware, system software
- 50 hrs. MTBI for applications
  - ~ 10^19 calculations w/o failure
- Light Weight Kernel compute node OS
- Highly redundant power supplies & memories
- Vertical blade and air flow design
- Independent RAS “immune” system
  - Processors, memory, network, power supplies, fans, disks
- “Over” design of connectors, fans etc.

# Economy

- Leverage commodity markets for components
  - Standard microprocessors, memory, Linux etc.
  - Customization only where truly critical to performance, e.g. network chip, LWK
- Design for lifecycle costs
  - Air cooled
  - Upgrade strategy
- Encourage a wider market for system
  - Amortize costs
  - Amplify impact

# Red Storm's future

- Sandia will upgrade to125TF this summer
  - Fifth row
  - Dual core processors
  - Enhanced network chip
- Red Storm is now Cray's primary product line
- 15 Red Storm/XT3 systems procured to date
  - NSF, DOD, Other DOE labs, Europe, Japan, …
- DOE Leadership Computing Facility (ORNL)
  - 100TF in 2006
  - 250TF in 2007
  - 1000TF in 2008