SAND2010-3413P

# Sandia's Computer and Information Sciences

## An overview for the External Review Panel

**Robert W. Leland, Director**
Computation, Computers, Information and Mathematics Center
May 26-28, 2010

**Outline:**

- **Changes in leadership**
- **Strategic perspective**
- **Technical and programmatic highlights**
- **Health measures for CIS**

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

# Substantial changes in CIS leadership and scope this past year

*Ted Blacker, Manager Computational Simulation Infrastructure (1543)*

*Rob Leland, Director Computation Computers, Information & Mathematics (Center 1400)*

*James Costa, Senior Manager Assured Comp & Info Security (8950)*

*Chuck Duus, Manager Cognitive Science Application 2 (1434)*

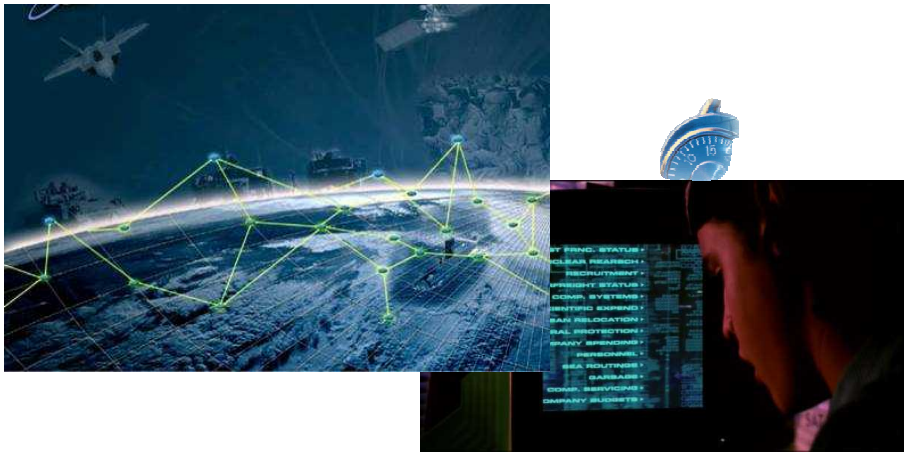
*John Wagner, Manager Cognitive Science Application 1 (1432)*

*Robert Hutchinson, Senior Manager Computer Sciences & Info Ops (8960)*

Sandia National Laboratories

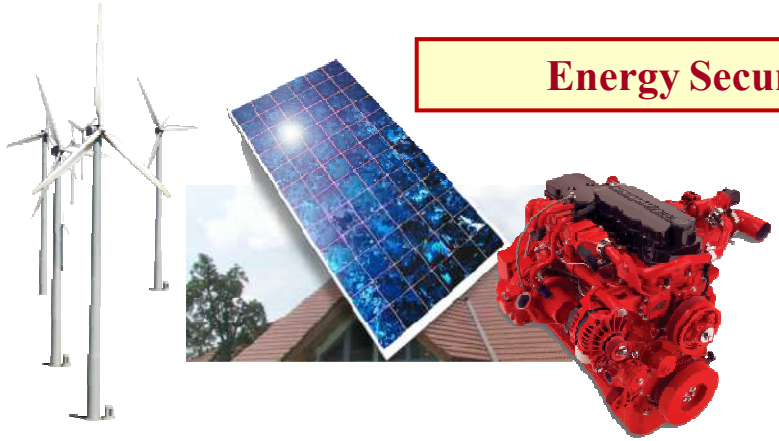
# Sandia has committed to three strategic thrusts

## Cyber Security



## Nuclear Security

## Energy Security

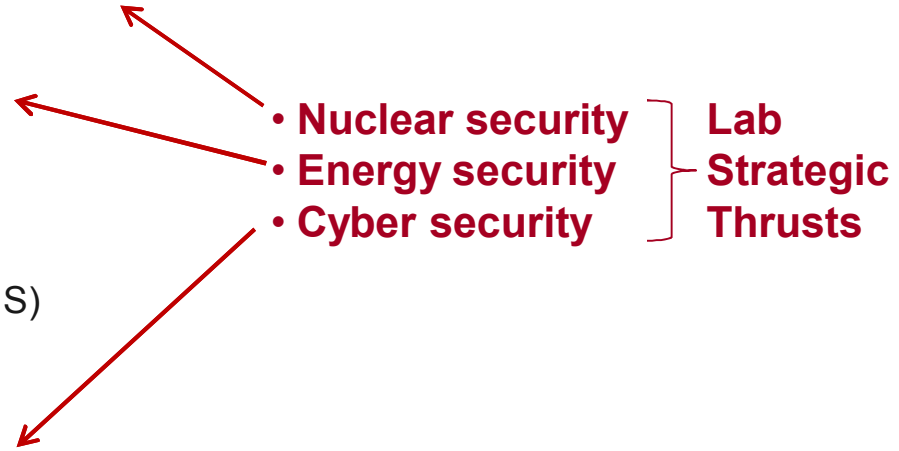




Sandia National Laboratories

# CIS will contribute primarily through four application themes with goals of high mission impact & creating new CIS missions

- **Nuclear weapons engineering**
  - Qualification Alternatives to Sandia Pulsed Reactor (QASPR)
  - Uncertainty Quantification
- **Energy/climate security**
  - Smart Grid
  - Nuclear Energy M&S
  - High fidelity climate model
  - Green House Gas Info. System (GHGIS)
  - Net Zero Carbon
  - Sub-surface modeling
  - Assessment and adaptation
- **Cyber assurance**
  - Trusted hardware and software stacks
  - Scalable informatics for cyber security, defense and attribution
  - Quantum computing as a game changer
- **Decision support**
  - "Certification" environments, e.g. NW or GHGIS
  - "High intensity" environments, e.g. combat teams or UAV management

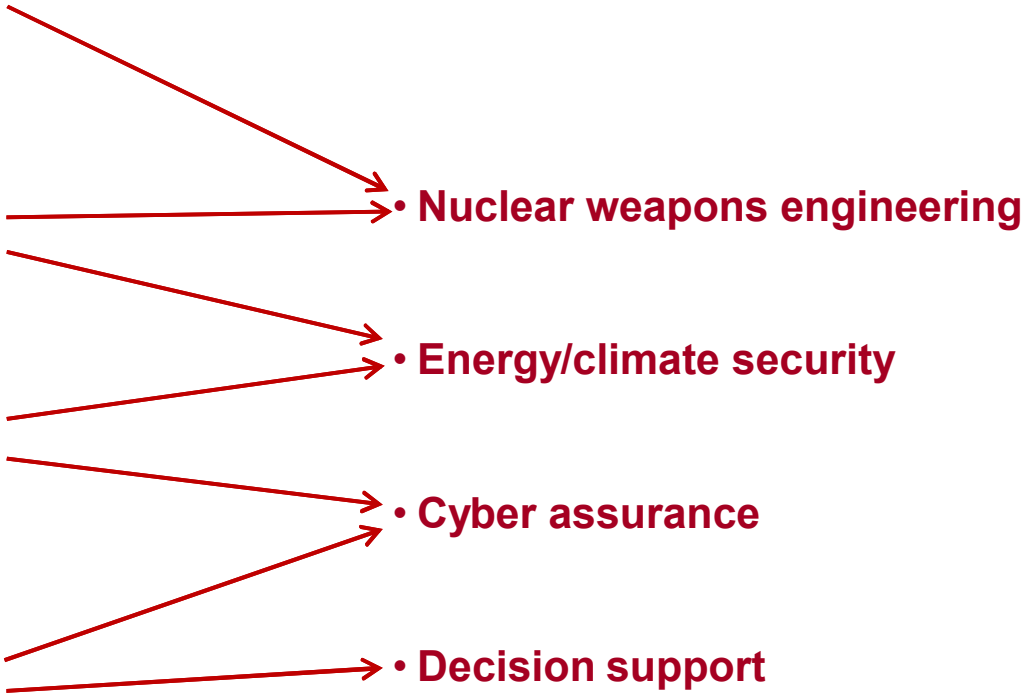





Sandia National Laboratories

# The major application themes leverage five principal technology thrusts

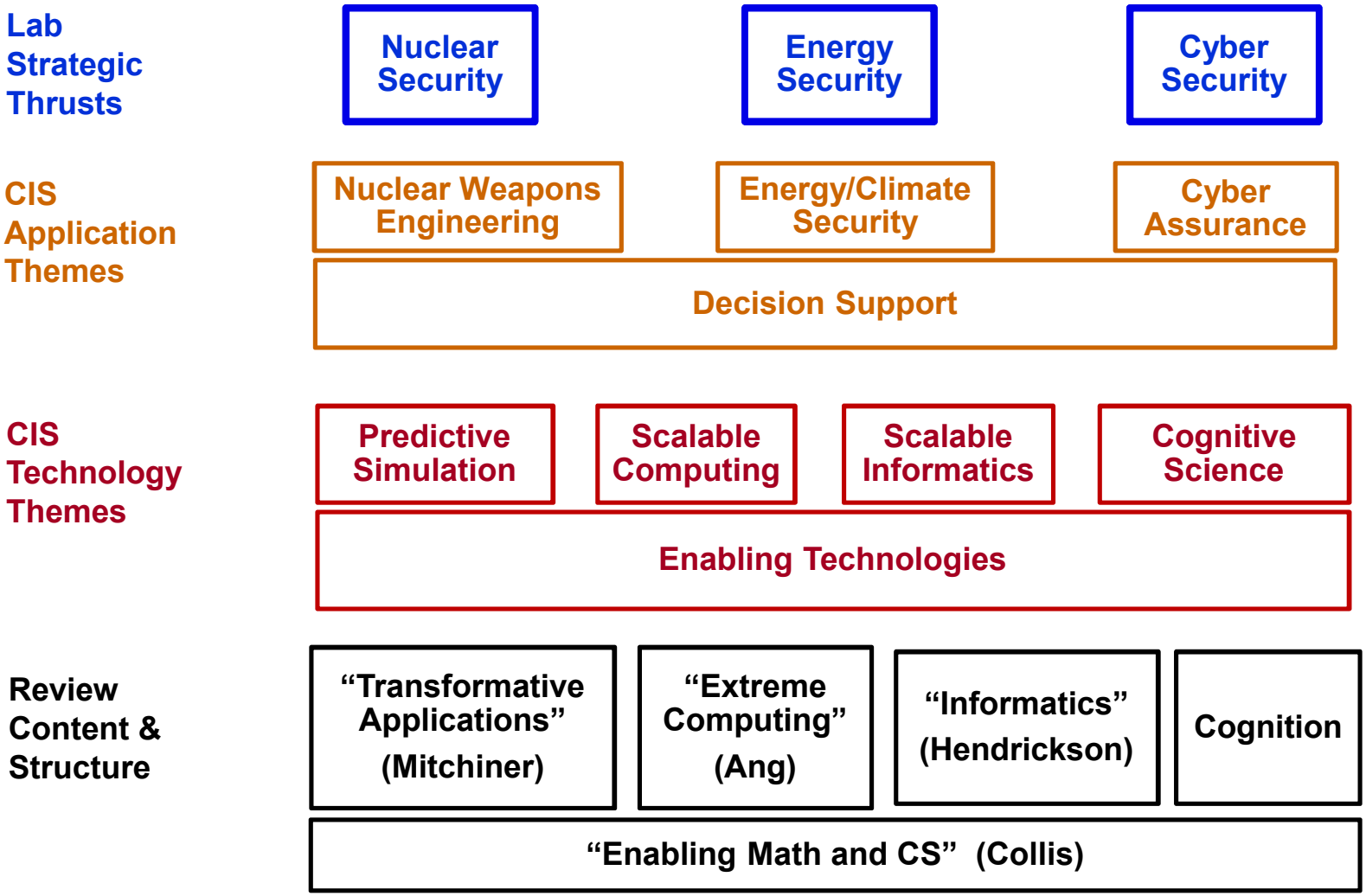
- **Extreme computing**
  - Exascale physical computing
  - Quantum computing
- **Predictive simulation**
  - UQ expertise applied to climate
- **Scalable informatics**
  - Vertically integrated effort
  - Diverse customer base
- **Cognitive sciences**
  - Engineering vs. sciences
- **Enabling technologies**
  - Mathematics & computer science






Sandia National Laboratories

# The CIS portfolio is strongly linked to Lab strategy






Sandia National Laboratories

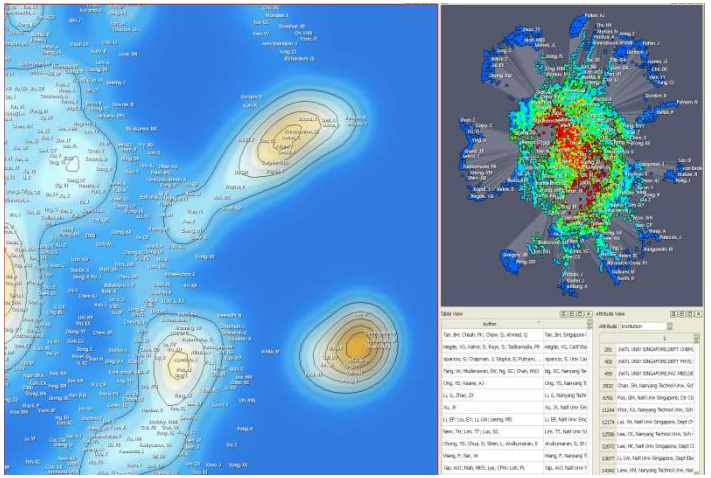
# Maximizing the CIS contribution will require finding the synergy between two potentially competing objectives

- Sustaining the historical depth, breadth and excellence of our research effort
- Finding ways to increase the actual and perceived mission impact of our work

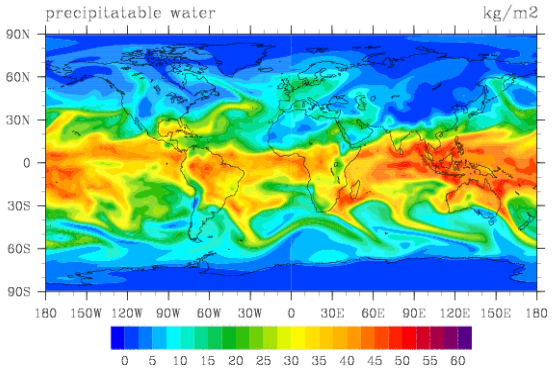
**…this is the strategic imperative for CIS**

# CIS is positioning for growth in new areas supporting new missions

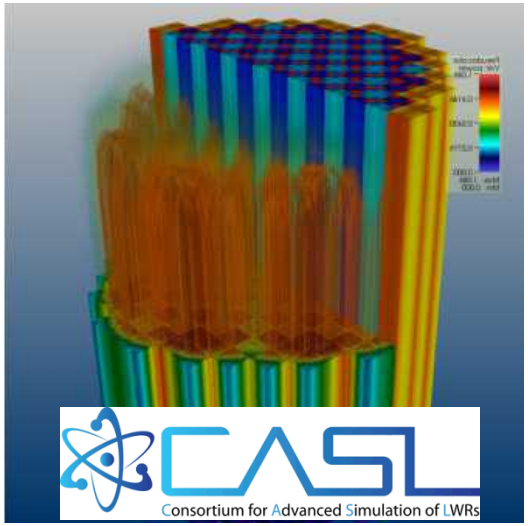

*GreenHouse Gas Information System (GHGIS) is Partnership (JPL/LANL/LLNL/SNL) to Provide Climate Change Monitoring*



*Key capability for Cyber business development*




*HOMME earth simulation with 100km model on 96k cores*




*SNL is leading member of the CASL hub proposal for advancing the design of nuclear reactors.*

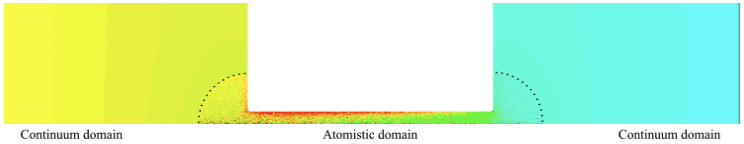


Sandia National Laboratories

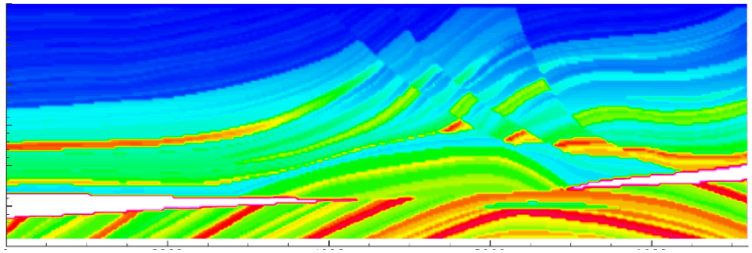
# Enabling technologies: a differentiating strength for CIS, generating strong growth and opening new markets

## Status

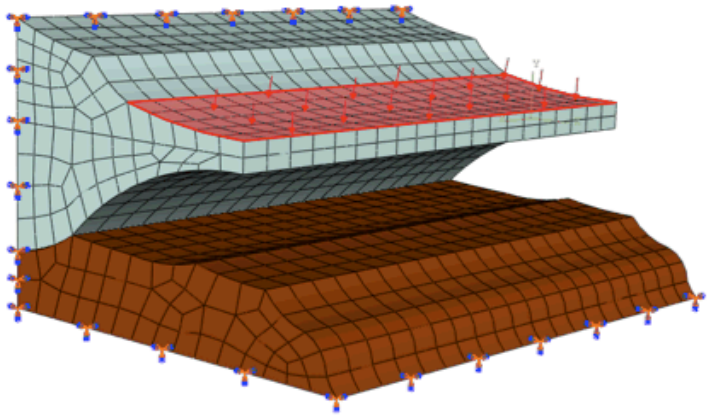
- **Significant growth in DOE ASCR**
  - 110% increase in funding since 2007
  - Multiscale, networks, UQ, optimization, climate, energy
- **New ASCR joint Math/CS Institute**
  - Architecture aware algorithms a central theme
- **Major new software releases**
  - Trilinos 10.x, DAKOTA 5.0, Acro 2.0
  - Software Engineering Maintenance Support team
- **Growth in architectures and GPU research**
  - Driven by subsurface modeling algorithms work
- **SciDAC-e projects in Carbon Sequestration**
  - builds on multi-grid and meshing R&D.
- **Growth in CSRI postdoc program**
  - 8 new postdocs since last review




**Multiscale Coupling**



**Subsurface Modeling**



**Ice-sheet Modeling**

Trilinos

DAKOTA

Acro

A Common Repository for Optimizers

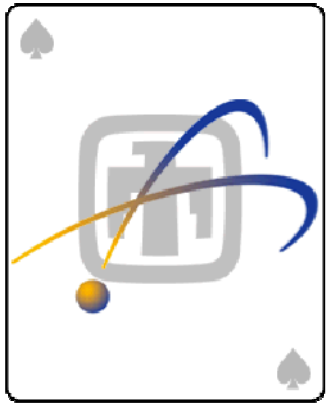


Sandia National Laboratories

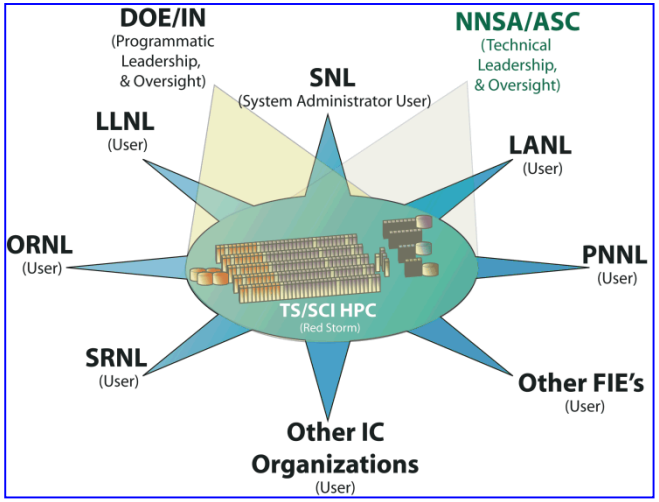
# CIS will continue as a national leader in high-performance computer architectures, system software, and algorithms

## Status

- **Red Sky design delivers Top 10 performance**
  - Mid range petascale system
- **Dark Storm supporting high impact apps**
  - Operations model under negotiation
- **Cray awarded ACES Cielo contract**
  - Sandia design influence evident
- **IAA projects moving forward robustly**
  - Awarded ASCR CS/Math Institute
- **Strong DARAP UHPC team assembled**
  - Status report at review
- **SNL central to SPEC Exascale alliance**
  - Co-design a key element of DOE strategy








*SRD/SCI DOE HPC Center*

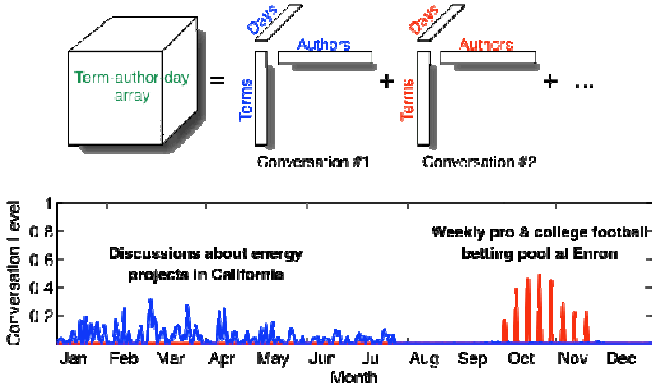


Sandia National Laboratories

# CIS intends to emerge as a national leader in algorithms, architectures and synthesis for scalable informatics

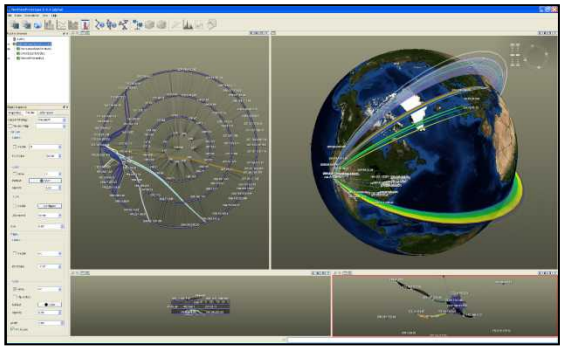
## Status

- **Focusing of effort**
  - Cyber security & defense
  - Decision support
- **Architectures**
  - XMT deployments
  - Netezza deployments
  - DARPA UHPC proposal
- **Collaborations**
  - Network Grand Challenge LDRD
  - Kachina (OGA)
  - Sandia cyber security & CI
  - Cray
  - Sun (Oracle)
  - Lexis/Nexis




*SNL Tensor Analysis of Enron Email Dataset*

*CRAY XMT - Multi-threaded HPC*



*Network GC Demos*



Sandia National Laboratories

# CIS success measures are being developed to support a strategic perspective

## CIS critical success factors

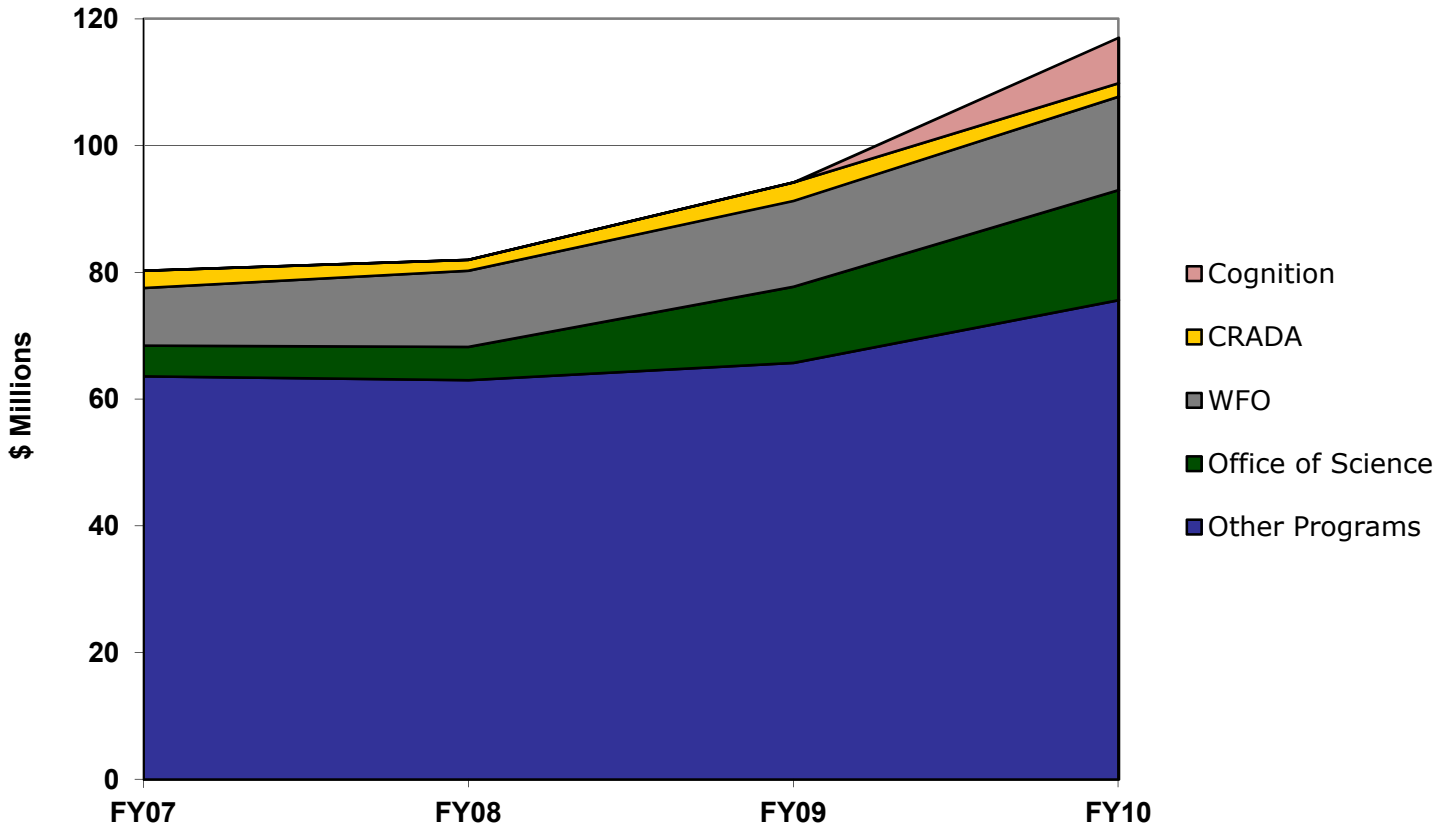
Target measurements

| World class workforce | Solving hardest problems | International recognition | Healthy investment & growth |
|---|---|---|---|
| ❑ Top quality staff<br>☑ Active student pipeline<br>❑ Targeted hiring<br>☑ Job satisfaction<br>☑ Employee retention<br>☑ Work fragmentation | ☑ ❑ Customer surveys<br>○ Satisfaction<br>○ Impact<br>❑ Internal surveys<br>○ Satisfaction<br>☑ ○ Milestone completion<br>☑ ❑ Citation volume | ☑ ❑ National/international presence:<br>○ Awards<br>○ Committees & editors<br>○ Keynotes, etc<br>❑ SW/HW Impact<br>☑ ○ Downloads<br>○ Usage<br>☑ ❑ Visitor volume | ☑ ❑ Revenue growth<br>☑ ❑ Success rate in capturing new business<br>❑ Internal investment in key areas |



Sandia National Laboratories

# CIS revenue is growing substantially in real terms

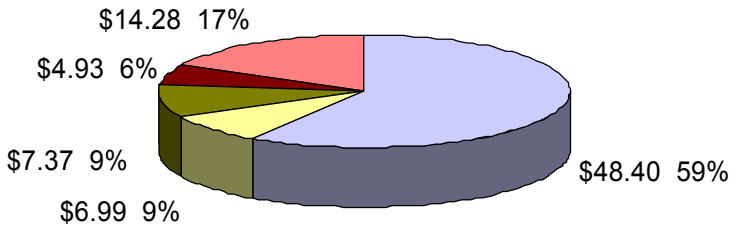




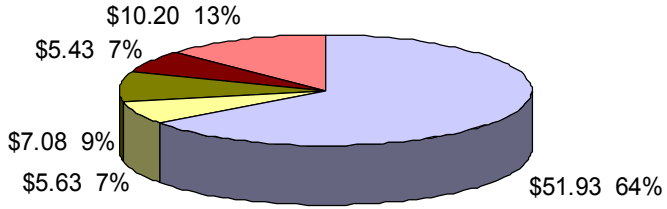

Sandia National Laboratories

# CIS budget changes over the past 4 years demonstrate substantial progress on mission diversification

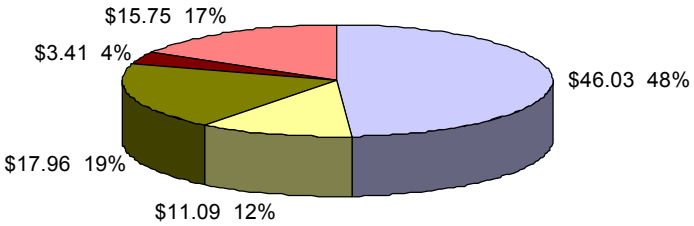
## FY07 ($80.3M)



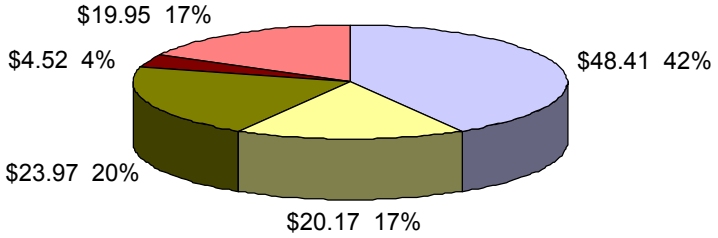

## FY08 ($82.0M)




## FY09 ($94.2M)




## FY10 ($117M)




NW DSA ERN HSD LDRD



Sandia National Laboratories

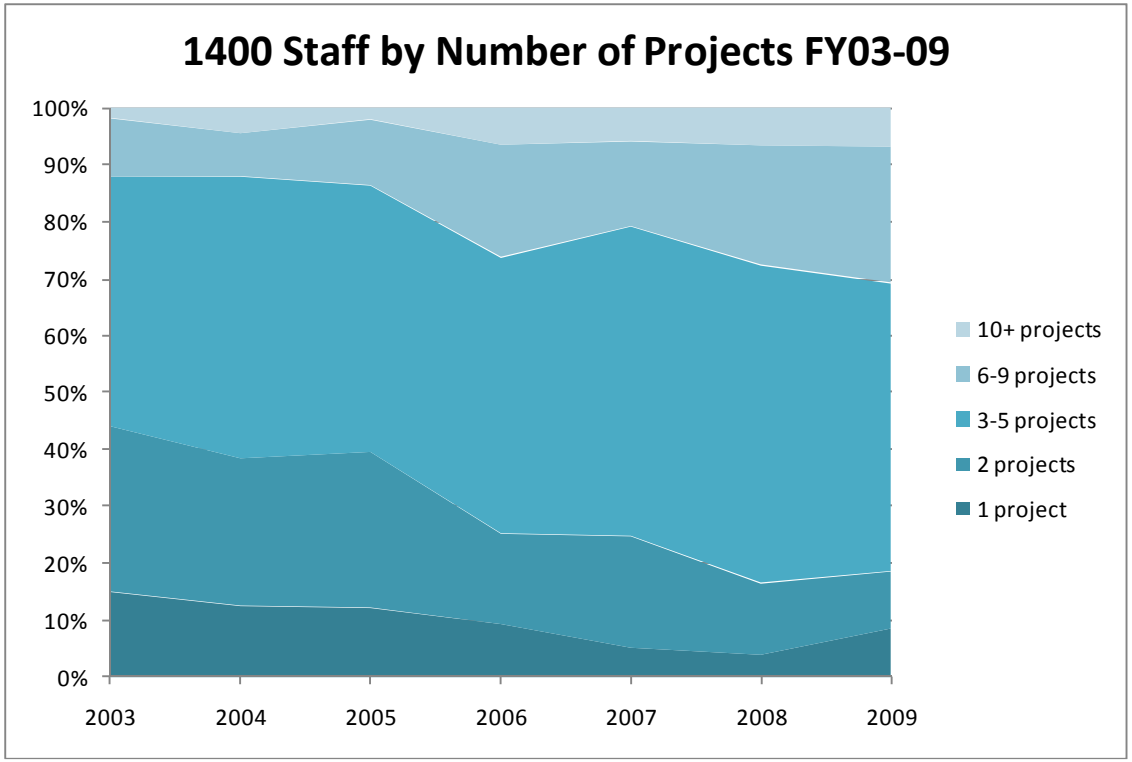
# With that diversification has come growing fragmentation

- Multi-tasking on several projects could mean less time for research because time is lost switching between tasks, and administrative burden accompanies each task.

## Current status and analysis:

Concern – Watch

- Half as many staff charged 1 or 2 projects last year than in 2003. Twice as many charged 6 or more projects.
- Mean number of projects charged by full time staff rose from 2.8 in 2003 to 4.8 in 2009.




Source: SNL Data Warehouse and STE SMU staff

Caution: Data includes all staff, not just technical staff, and excludes Group 8960

# Sandia's CIS paper impact record is very competitive with the other NNSA Labs

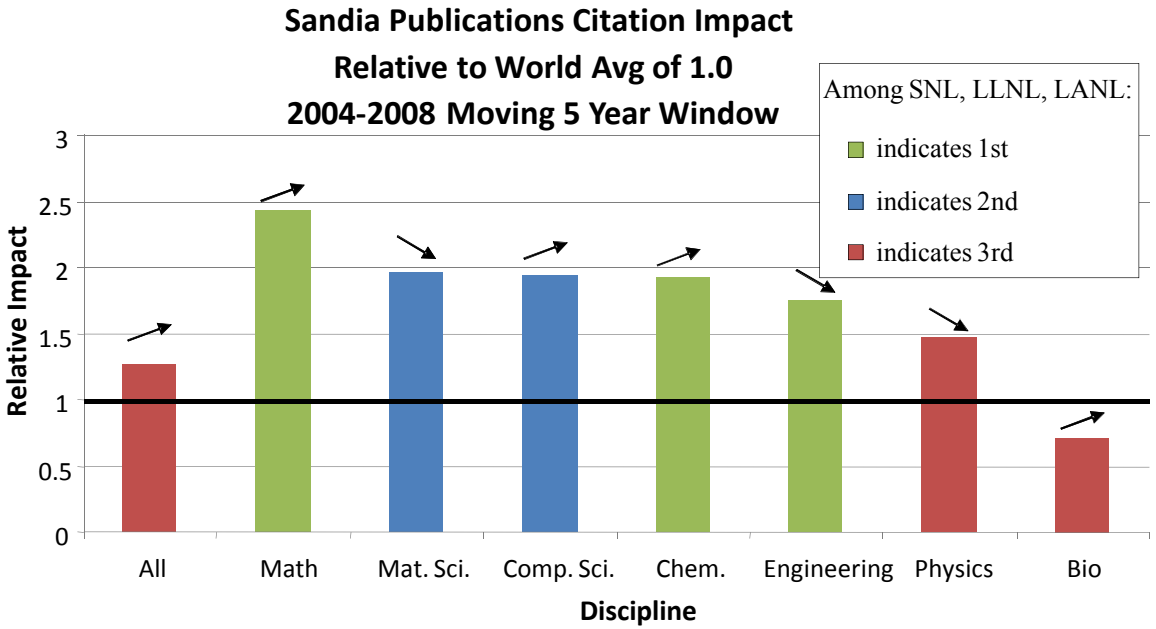
## Measured because:

- Citations indicate that our publications have been read and used. Thomson Reuters calculates citation impact compared to a world average in that research area.

## Current status and analysis:

Meets expectations

- Sandia is above the world average in number of citations per paper (compared within fields) with exception of Biosciences, a new area for us.
- In Computer Science SNL is 1.94 to LLNL's 2.04 and LANL's 1.78. SNL published 180 papers to LLNL's 145 and LANL's 190.
- In Math SNL is 2.43 compared to 2.13 for LLNL, 1.93 for LANL. SNL published 52 papers compared to LLNL's 41, LANL's 121.




Arrows show trend since 1990s.

**N B: CIS researcher publish across all these topics; Math and CS combined account for < 40% of CIS pubs.**

Source: Thomson Reuters, 2008 data purchased in April/May 2009, and Sandia Technical Library staff



Sandia National Laboratories

# Sandia's paper impact and volume in computer science is competitive with a wider set of major labs

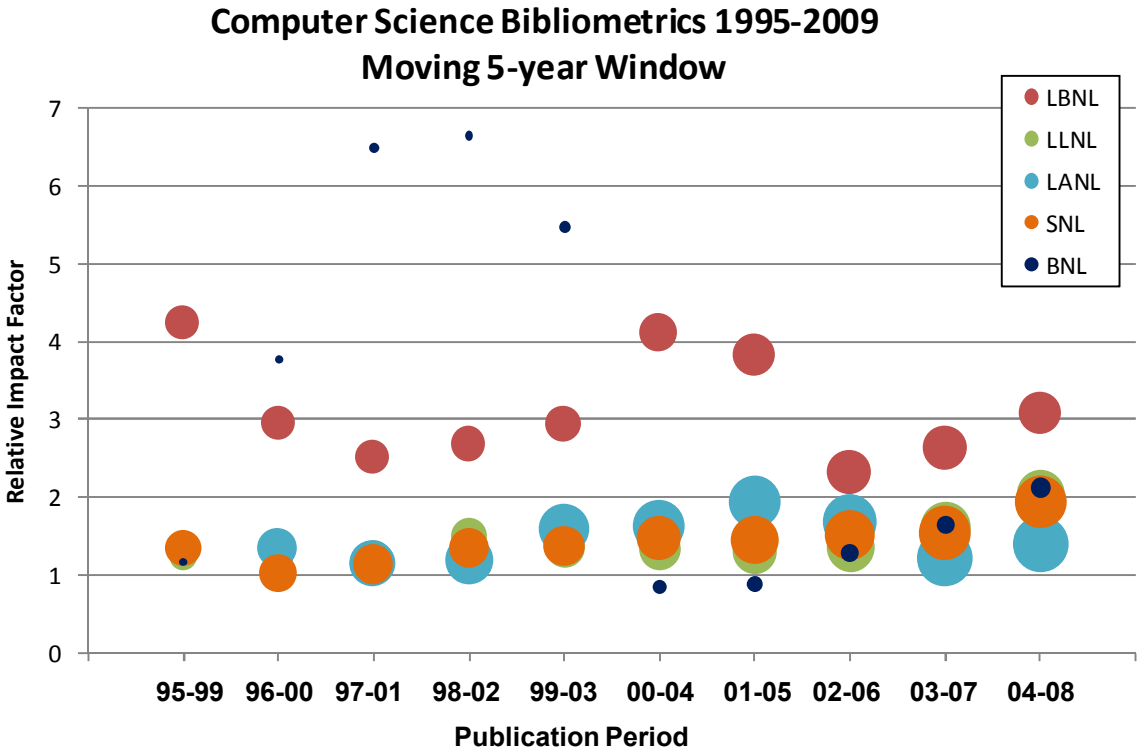
## Measured because:

- Average citations per paper is influenced by total number of papers published.
- Comparing against others helps gauge our performance. Sandia purchases comparisons with 14 Laboratories.

## Current status and analysis:

Targets not set

- Sandia is 4th highest of the 14 labs in citation impact (1.94). LBNL is highest at 3.09 from 113 papers, then BNL at 2.13 and LLNL at 2.04.
- Sandia is 3rd highest in number of publications (180) behind NASA with 359 and LANL with 190.
- The number of Sandia CS publications has steadily risen from 80 in CY 95-99 to 180 in 04-08.
- LDRD is 33 of the 180 papers and has citation impact of 2.55 in the 04-08 period.



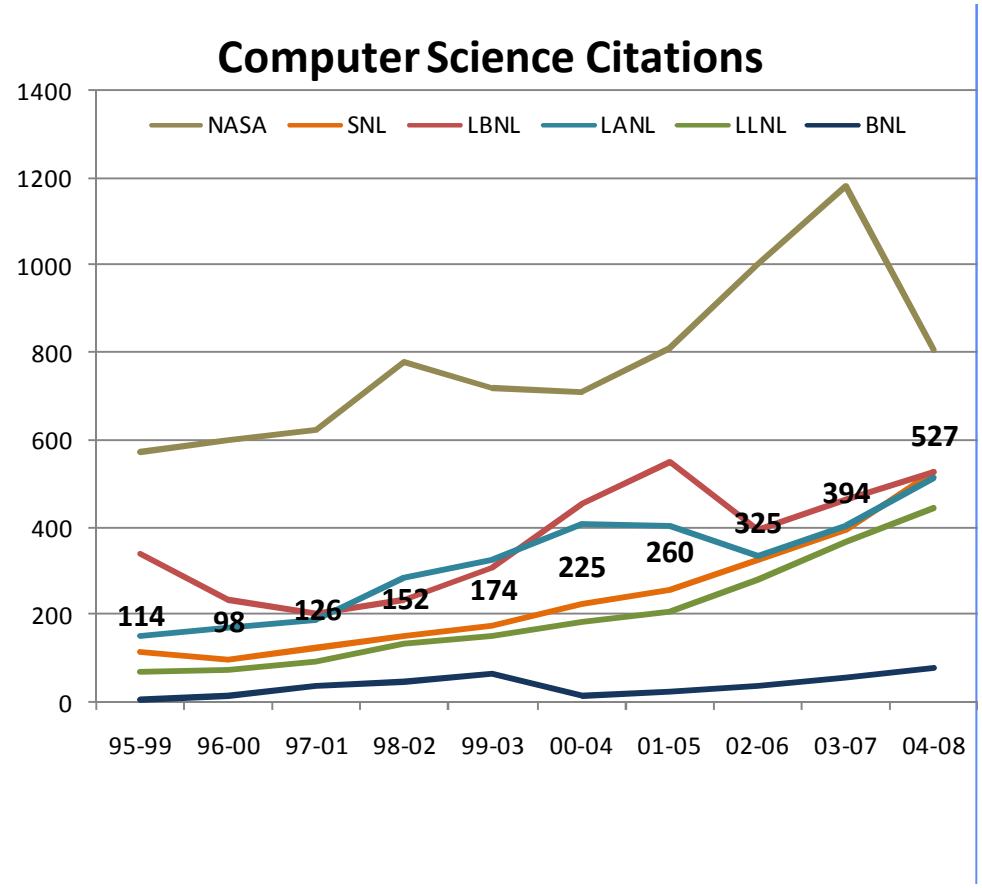

- Analysis covers SNL, LANL, LLNL and labs in the comparison set that had higher citation impact than SNL in this discipline.
- The larger the circle, the higher the number of publications.
- These are 5 year moving averages, so a highly cited paper may create a cliff when it falls out of the window.

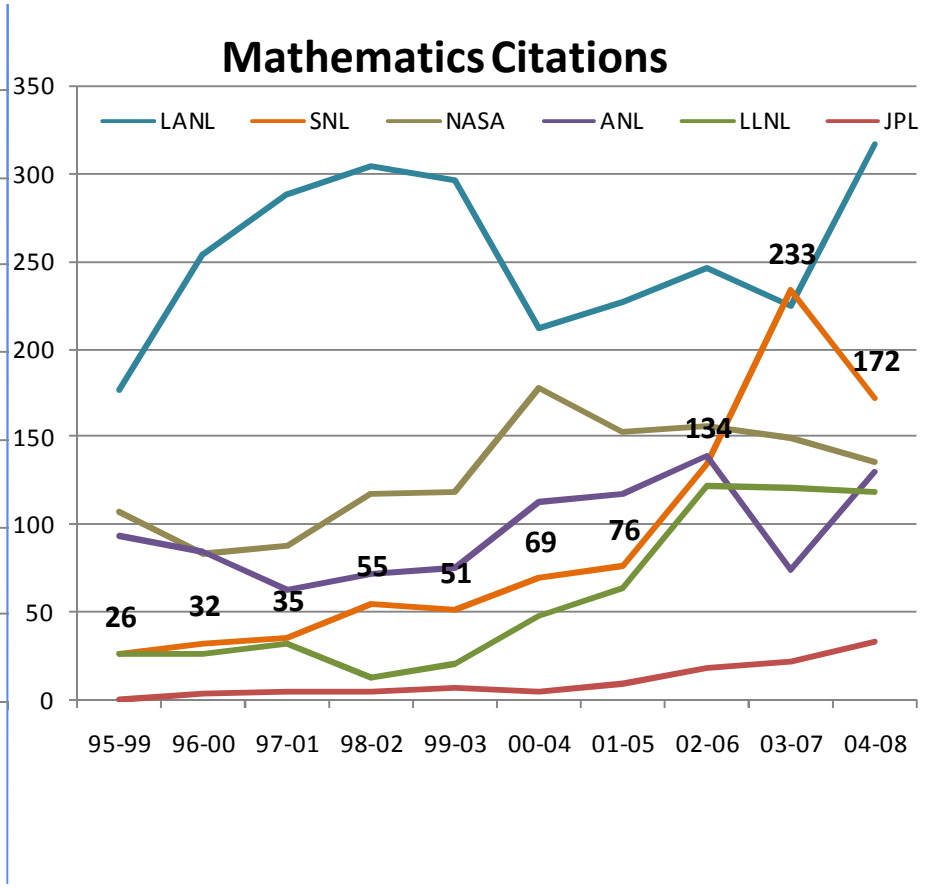
Source: Thomson Reuters and STE SMU staff



Sandia National Laboratories

# Sandia's total citation record in math and computer science is competitive with other major Labs




Source: Thomson Reuters and STE SMU staff



Sandia National Laboratories

# CIS is an applied research and development organization, which helps to calibrate our publication profile

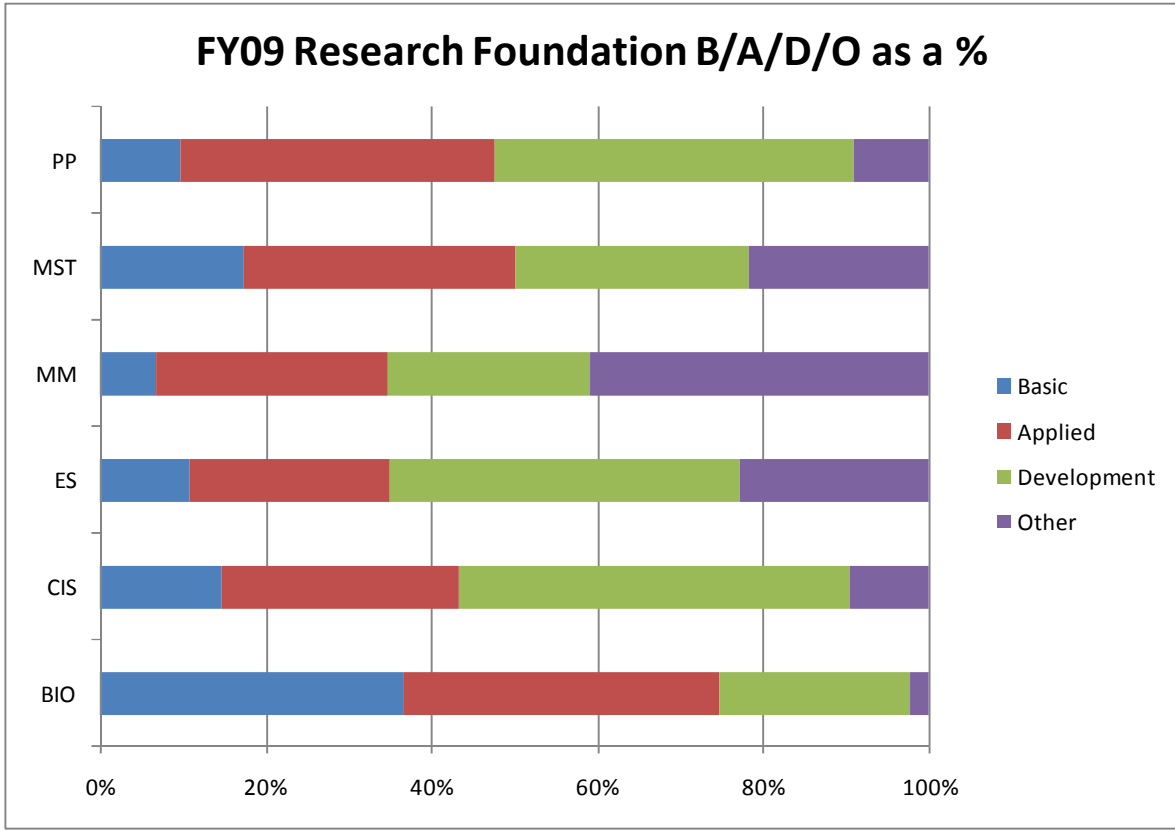
## Measured because:

- This mix is an indicator of who our customers are and to what extent SNL does basic and applied research.

## Current status and analysis:

**Meets expectations**

- In FY09, Basic research in the RFs was $80 M and 11% of total ST&E SMU expenditures, compared to $70 M and 12% in FY08.
- CIS RF mix of B-A-D was stable between FY 08 and FY 09.
- The same was true for the rest, other than Pulsed Power's portfolio, which shifted significantly.



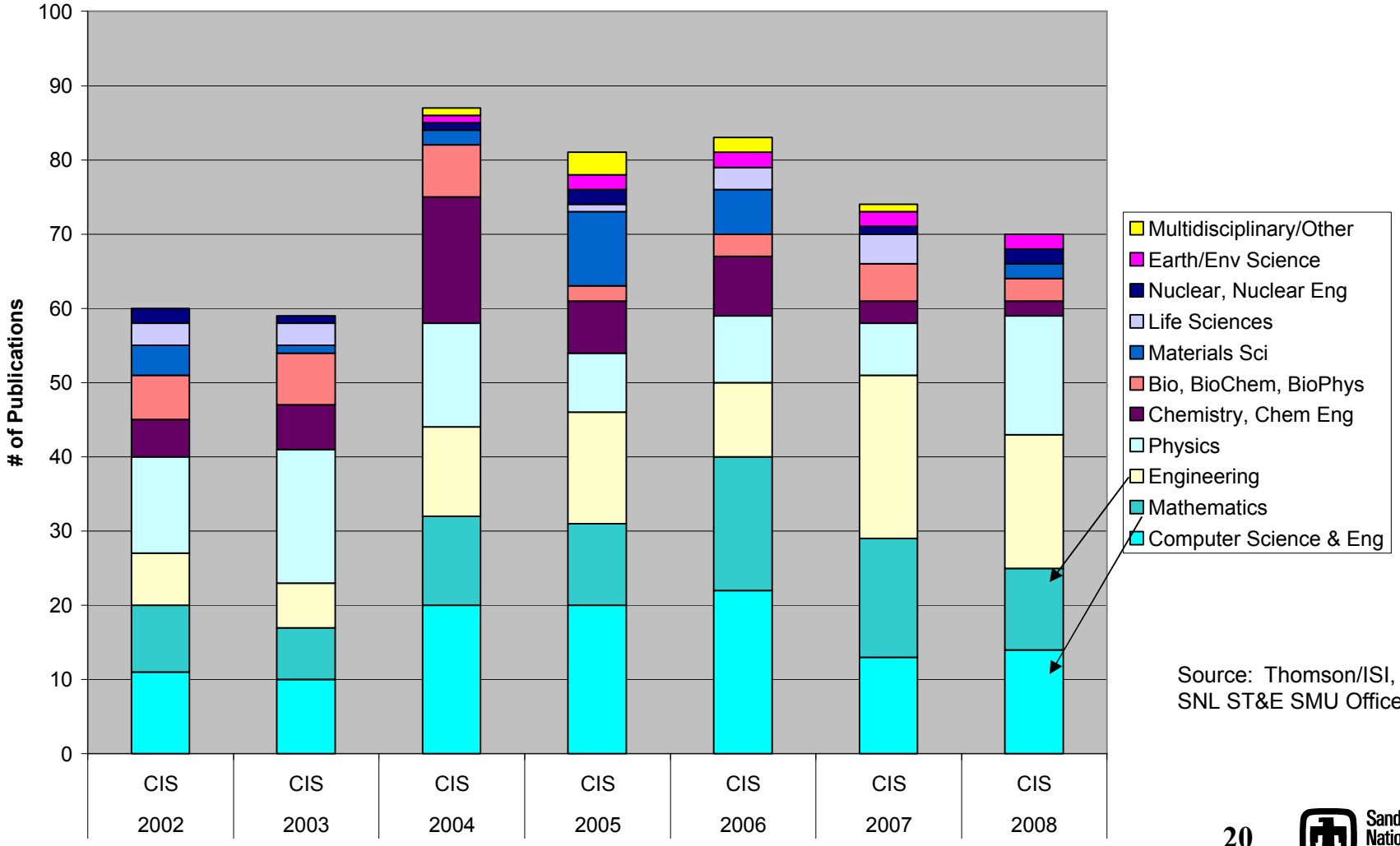

Source: internal analysis based on program manager judgments, and financial data from MRSO Reportville



Sandia National Laboratories

# CIS researchers publish across a diverse range of fields, hence math & CS papers capture only a fraction of the impact

**CIS Peer-Reviewed Publications by Discipline Area 2002-2008**




Source: Thomson/ISI, and SNL ST&E SMU Office



Sandia National Laboratories

# CIS presence in the technical community is significant

## Software Downloads (May 2009 – May 2010)

- Paraview – 121,217
- LAMMPS – 8,900
- Pyutilib – 8,469
- Jess – 8000
- Trilinos – 4,145
- DAKOTA – 3,419
- Coopr – 2,593
- Tensor Toolbox – 800
- OPT ++ – 360
- Zoltan – 356
- JAGUAR – 306 (12/09-4/10)
- Chaco – 300
- Acro – 246
- Sisyphus – 156
- CGM – 154
- Kitten – 143
- Poblano – 136 (rel. 3/10)
- FAST – 118
- VERDICT – 83
- TEVA-SPOT – 76
- LOCA 1.0 – 50
- Utilib – 49
- SMB – 40
- CIT – 40
- OpenCatamount – 34
- ChISELS – 31
- Facetbool – 22
- CPA2 – 20
- CANARY – 14
- App_model – 8
- showmesh – 6

**TOTAL: 160,291**

## CIS National/International Presence

- 2010 Award for Excellence in Technology Transfer (Red Storm)
- 2009 R&D 100 Award for Catamount N-Way Lightweight Kernel
- Patents (6)
- Books/Proceedings (6)
- Best Papers (1)
- Keynotes (4)
- Journal Editors (22)
- Conference Committees (59)
- Professional Society Promotion (3)

# We have substantial external engagement through CSRI

| Measure | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|
| Projects | 14 | 3 | 1 | 1 |
| Workshops | 4 | 7 | 5 | 2 |
| Visitors (Institutions) | 102(67) | 106(68) | 135(75) | 81(38) |
| Summer Students | 43 | 34 | 41 | 51 |
| Sabbaticals | 3 | 2 | 2 | 3 |
| % Hired having CSRI Experience | 73% | 80% | 75% | 33% |




Sandia National Laboratories

# CIS contributes substantially in other dimensions, e.g. through its influence on the HPC marketplace

*" Without a doubt, Cray's partnership with Sandia has been one of the most significant partnerships in the company's history. Not only did our collaboration on Red Storm result in significant scientific advancements and accomplishments in support of the Department of Energy's NNSA mission, it also spawned a very successful, proven line of Cray XT supercomputers starting with our XT3 system launched in 2005. What is even more significant is that the benefits of our relationship have extended beyond just Cray, Sandia and the NNSA. Numerous high performance computing centers around the world have turned to the XT's MPP supercomputing architecture to help solve some of the world's most pressing scientific challenges, from systems of a single cabinet to the largest supercomputer in the world (and the first system to run real scientific applications at a performance level of over a petaflops). All of this would not have been possible without the unique combination of talent and dedication that our partnership enabled. We are very grateful for Sandia placing a big bet on Cray, one that's paid off for both of us and the worldwide supercomputing community at large. "*

*Peter Ungaro, President and CEO, Cray*

*(2010 statement)*

# Supplementary material

# We are prototyping a software maintenance model for CIS

- **At the 1410 group level we have formed the SEMS Team (Software Engineering Maintenance and Support)**
  - Supports Trilinos, Dakota, ACRO and others
  - Includes sys admin, support and usability specialists, scientific programmers, SQE specialists
  - Status:
    - Detailed draft plan produced
    - Seeking SEMS project lead
    - Key Trilinos contractor converted to staff
    - Full-time user support lead for DAKOTA hired & matrixed
    - Seeking SEMS coordinator to replace Ken Alvin
- **1410 has formed a Software Licensing Working Group**



Sandia National Laboratories

# Sandia's paper impact and volume in mathematics is competitive with a wider set of major labs

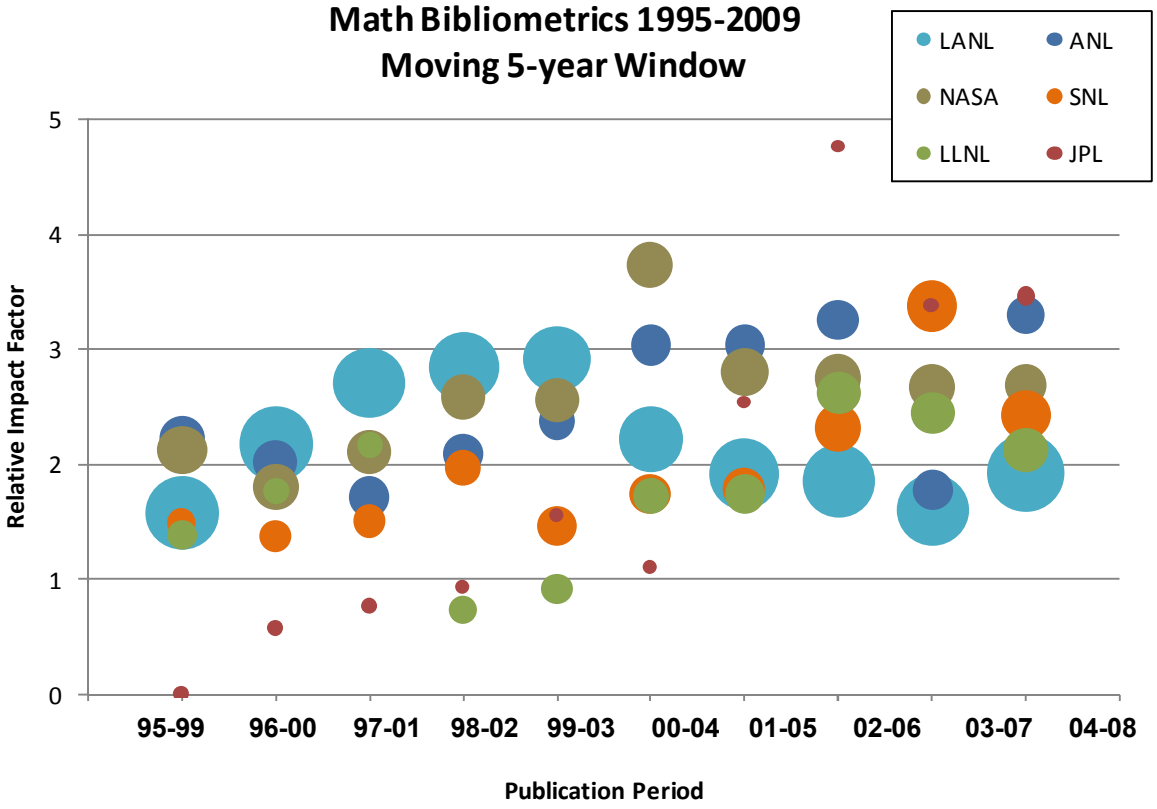
## Measured because:

- Average citations per paper is influenced by total number of papers published.
- Comparing against others helps gauge our performance. Sandia purchases comparisons with 14 Laboratories.

## Current status and analysis:

Targets not set

- Sandia is 4th highest of the 14 labs in citation impact (2.43). JPL is highest at 3.46 from 7 papers. ANL is 3.29, NASA 2.68.
- Sandia is 3rd highest in number of publications (52) behind LANL with 121 and NRL with 58.
- The number of Sandia Math publications has steadily risen from 17 in CY 95-99 to 52 in 04-08.
- LDRD is 8 of the 52 papers and has citation impact of 3.13 in the 04-08 period.

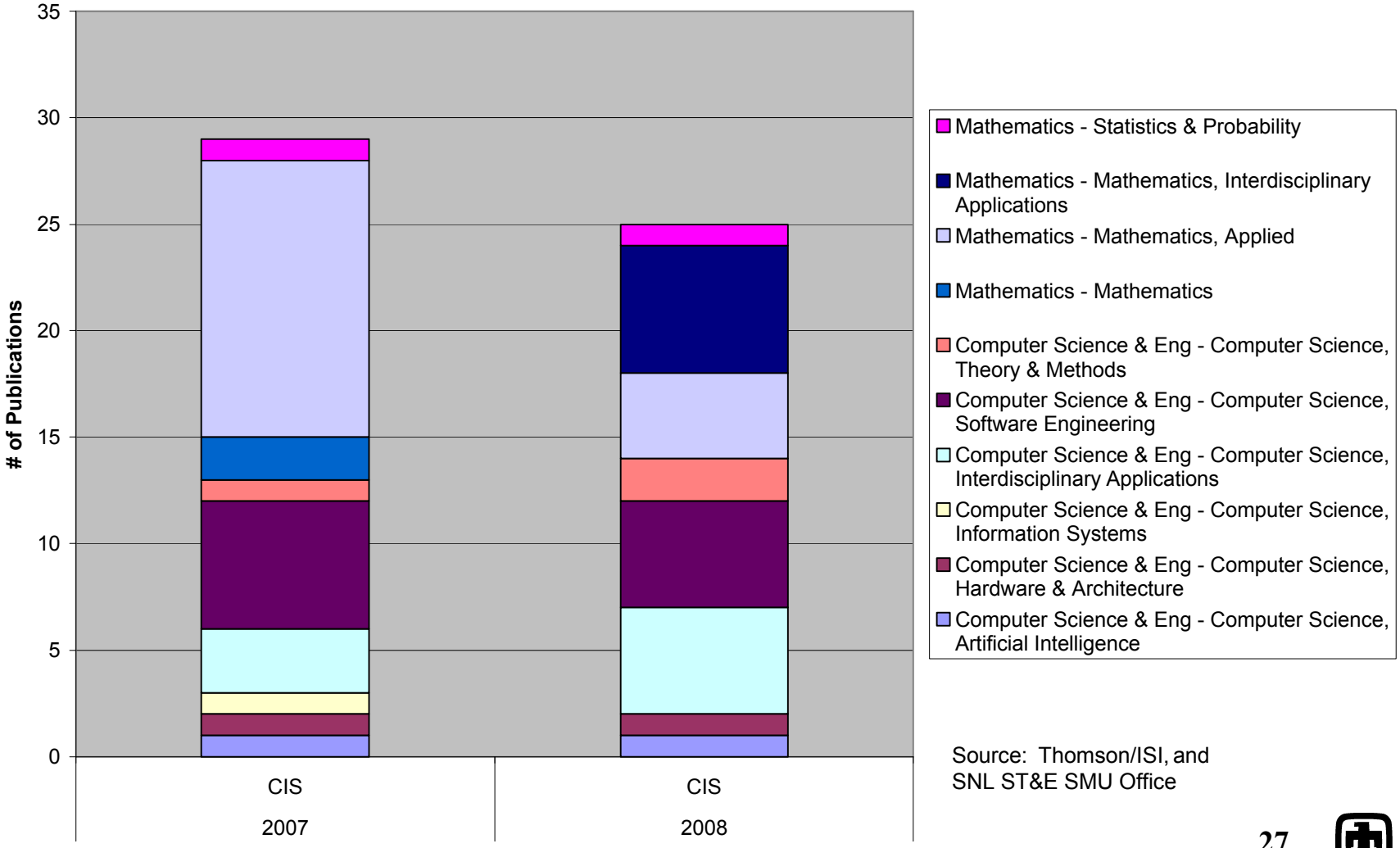



- Analysis covers SNL, LANL, LLNL and labs in the purchased comparison set that had higher citation impact than SNL in this discipline.
- The larger the circle, the higher the number of publications.
- These are 5 year moving averages, so if a highly cited paper is no longer in the set, average citations may drop precipitously if no others take its place.

Source: Thomson Reuters and STE SMU staff

# CIS publications cover a wide spectrum within math and computer science

**CIS Computer Science and Math detail 2007-2008**




Source: Thomson/ISI, and SNL ST&E SMU Office



Sandia National Laboratories

# Staff retention may be a moderate concern for CIS

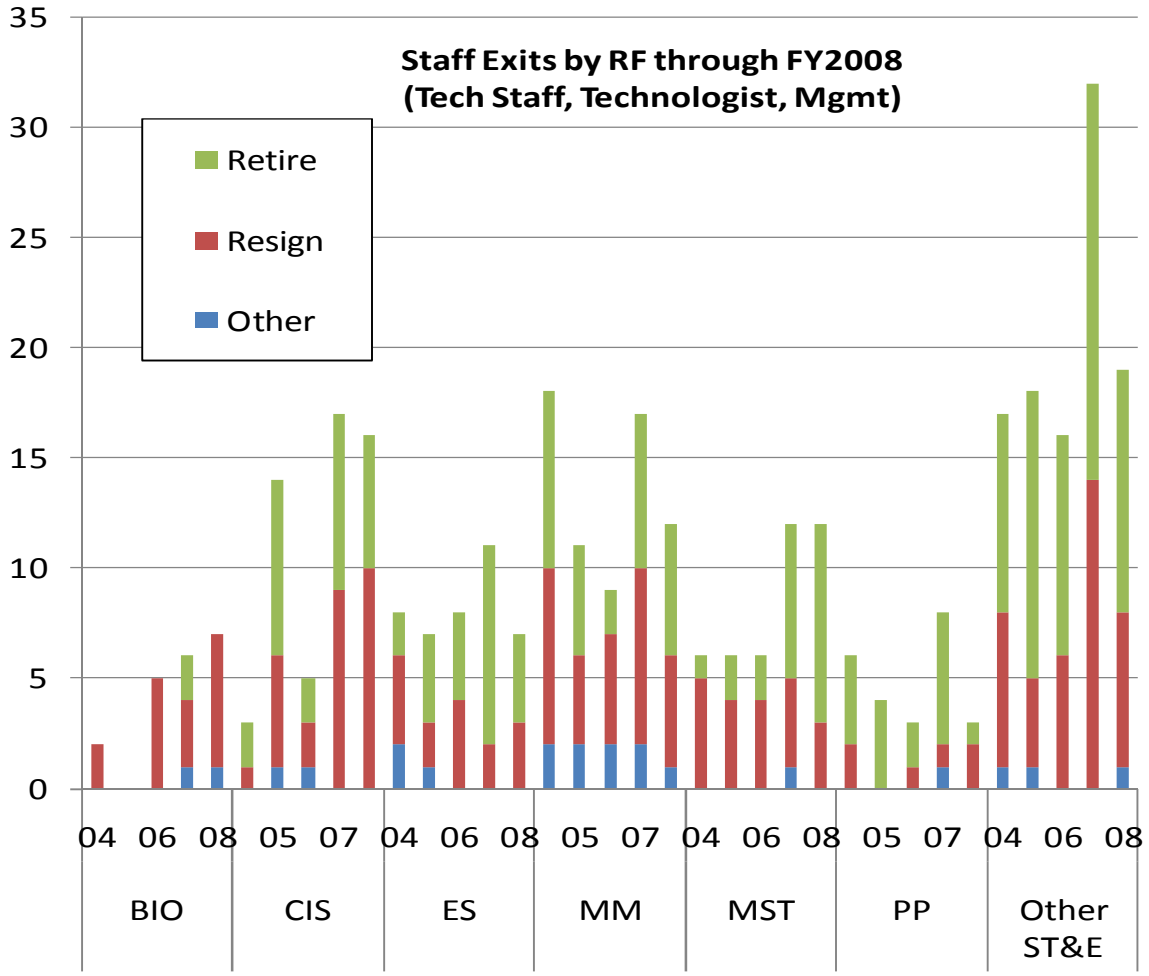
## Measured because:

- Retaining good staff is important. Number of exits is a leading indicator of future problems in getting the work done.

## Current status and analysis:

Meets expectations

- Exits from the ST&E SMU are about 4 percent of the STE SMU workforce
- Exits from the Lab include both those that resign and retire

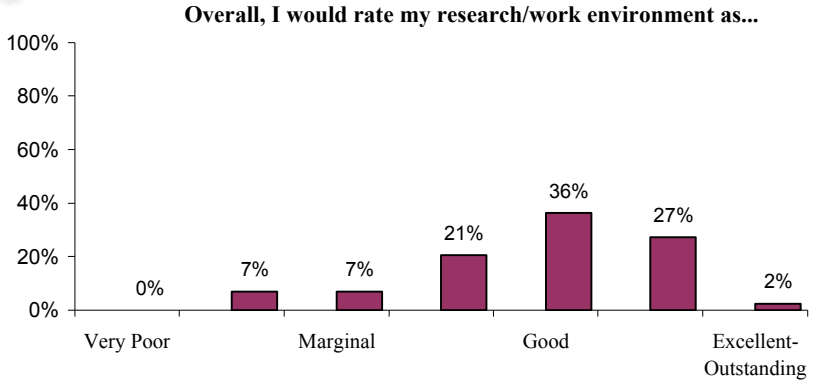


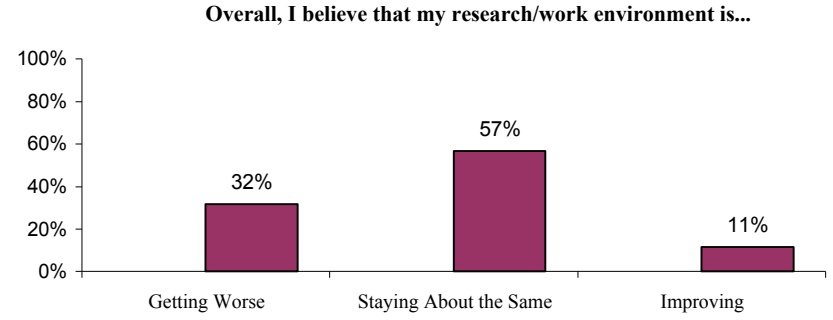

Source: Sandia HRIS

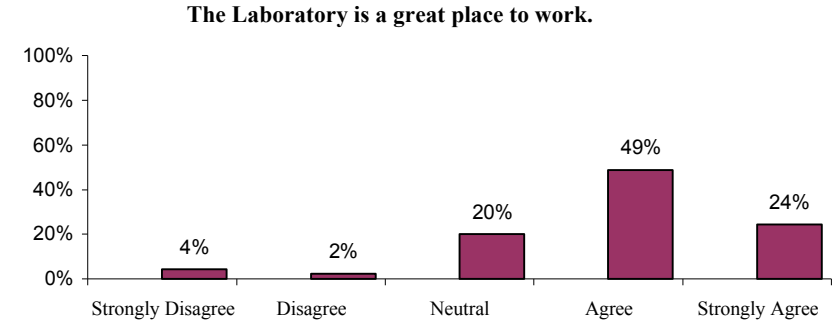


Sandia National Laboratories

# Health of CIS research environment is well above average for ST&E at Sandia




Sample Size - 74

- **CIS is happier**
  - Overall environmental rating is 65% good or better vs. 55% background
- **Particularly in these dimensions:**
  - Equipment,
  - Salaries,
  - Respect,
  - Stability of funding,
  - Research competencies,
  - Recognition of merit,
  - Career advancement,
  - Allocation of internal research funds,
  - Investment in new areas & basic research.
- **Overhead is measurably lower:**
  - Fewer projects worked
  - Less time spent on compliance;
  - More conferences attended

# CIS external customer satisfaction is high
## (6 projects within Sandia corporate sample)

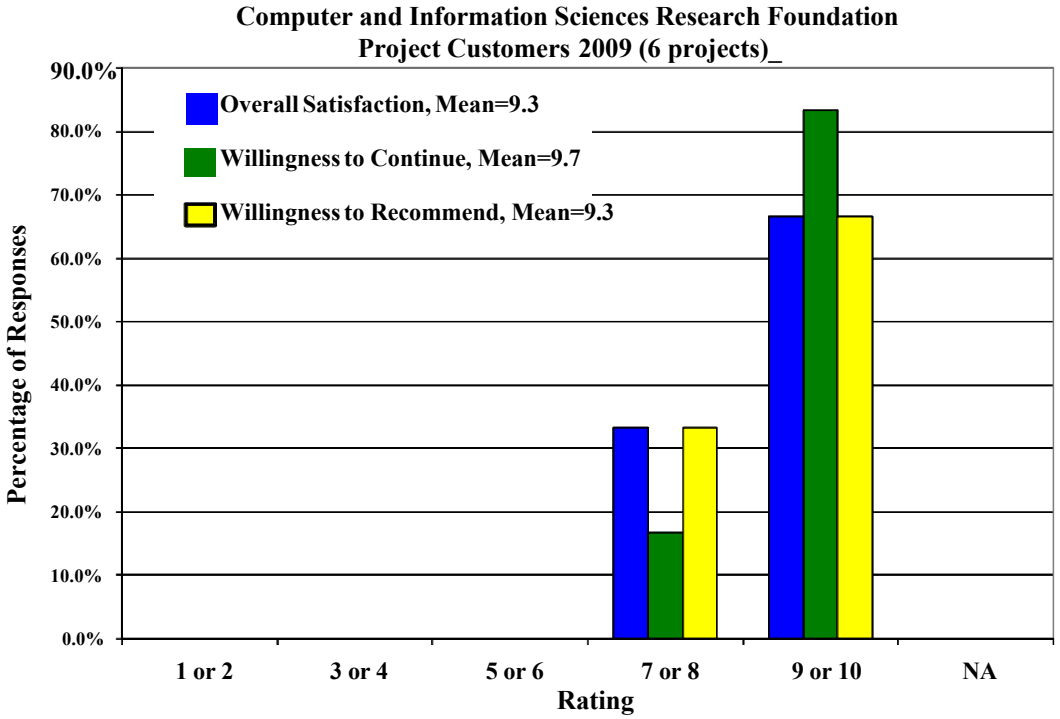
### Measured because:

- Feedback from our external customers tells us how we are doing and suggests where improvement is needed.

### Current status and analysis:

Meets expectations

- Overall satisfaction of 9.3 on a scale of 10 is very good. Sandia corporate average was 9.0. CIS was 8.7 in 2008.
- Detailed questions showed "5" on scale of 5 in areas of Staff, Relationship, Diversity, and Communication.
- An area of possible concern because it rates lower and customers see it as important is Schedule, including meeting scheduled milestones (down to 3.6 from 4.5 in 2008).



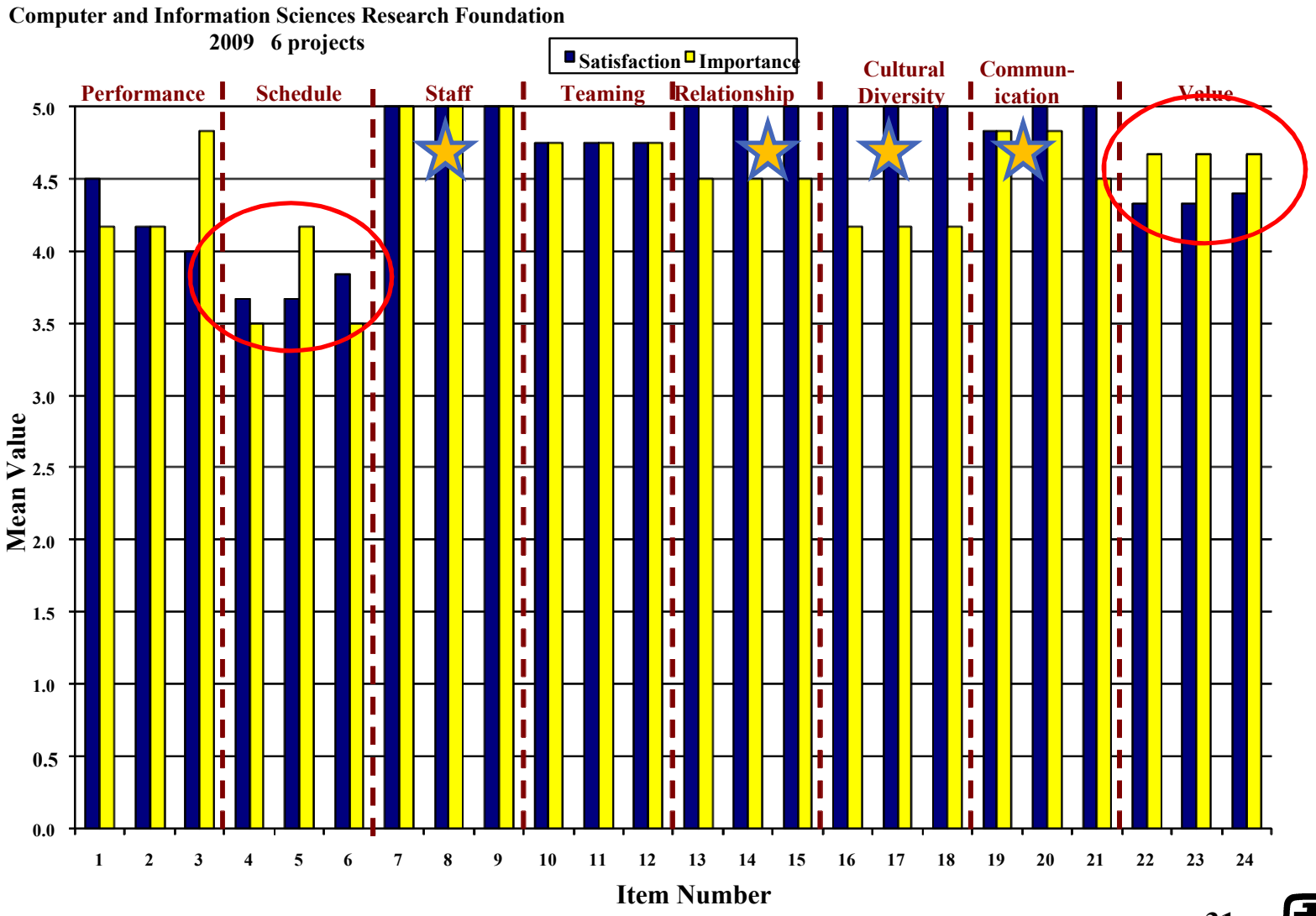

Source: Sandia corporate CSAT, Dept. 9751

Caution: These 6 projects were chosen at random and may not represent the CIS RF as a whole.



Sandia National Laboratories

# CIS external customer satisfaction compares well with the rest of the Sandia ST&E community






Sandia National Laboratories

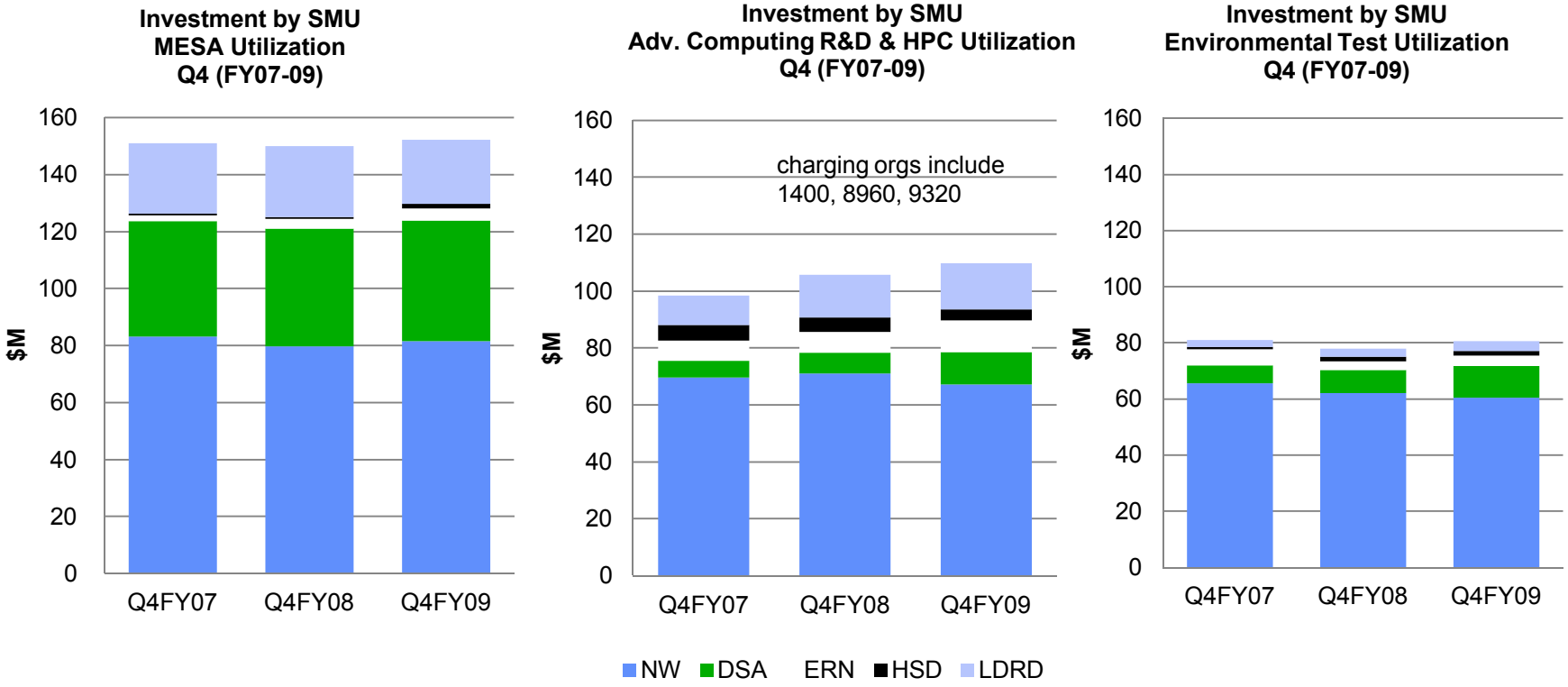
# SNL internal investment in HPC is increasing moderately

**Measured because:**

**The Capabilities Forum is interested in investment trends across SMUs.**

**Current status and analysis:**

**Concern -Watch**




Source: financial data from MRSO Reportville

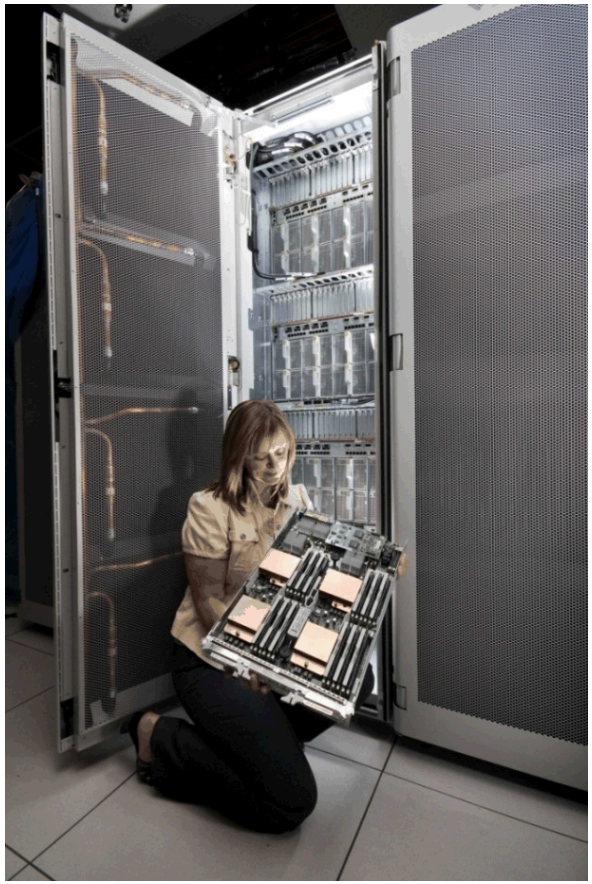


Sandia National Laboratories

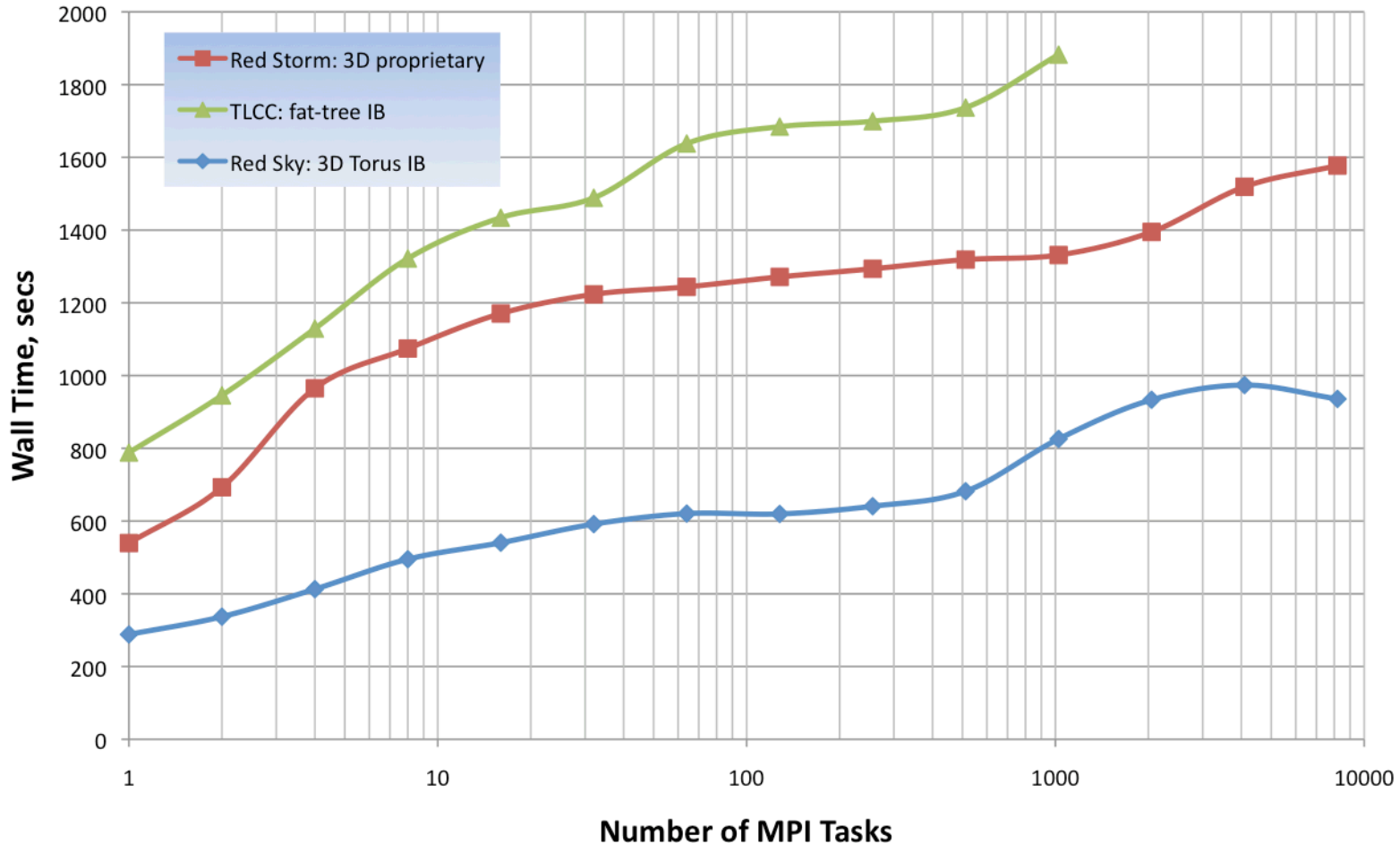
# Red Sky demonstrates a large number of significant technical innovations

- **Bridging from capacity to capability**
  - Many “Red” characteristics at commodity price
  - 2-3X faster than Red Storm in mid range
  - 1/3 operational costs

- **Top ten Red Sky innovations**
  - Petascale midrange system
  - Intel Nehalem processor
  - QDR Infiniband
  - 3D mesh/torus
  - Optical cabling
  - Optical Red/Black switching
  - Refrigerant cooling/ glacier doors
  - Power distribution
  - Interconnect resiliency
  - Minimal ethernet (RAS & mgmt. only)



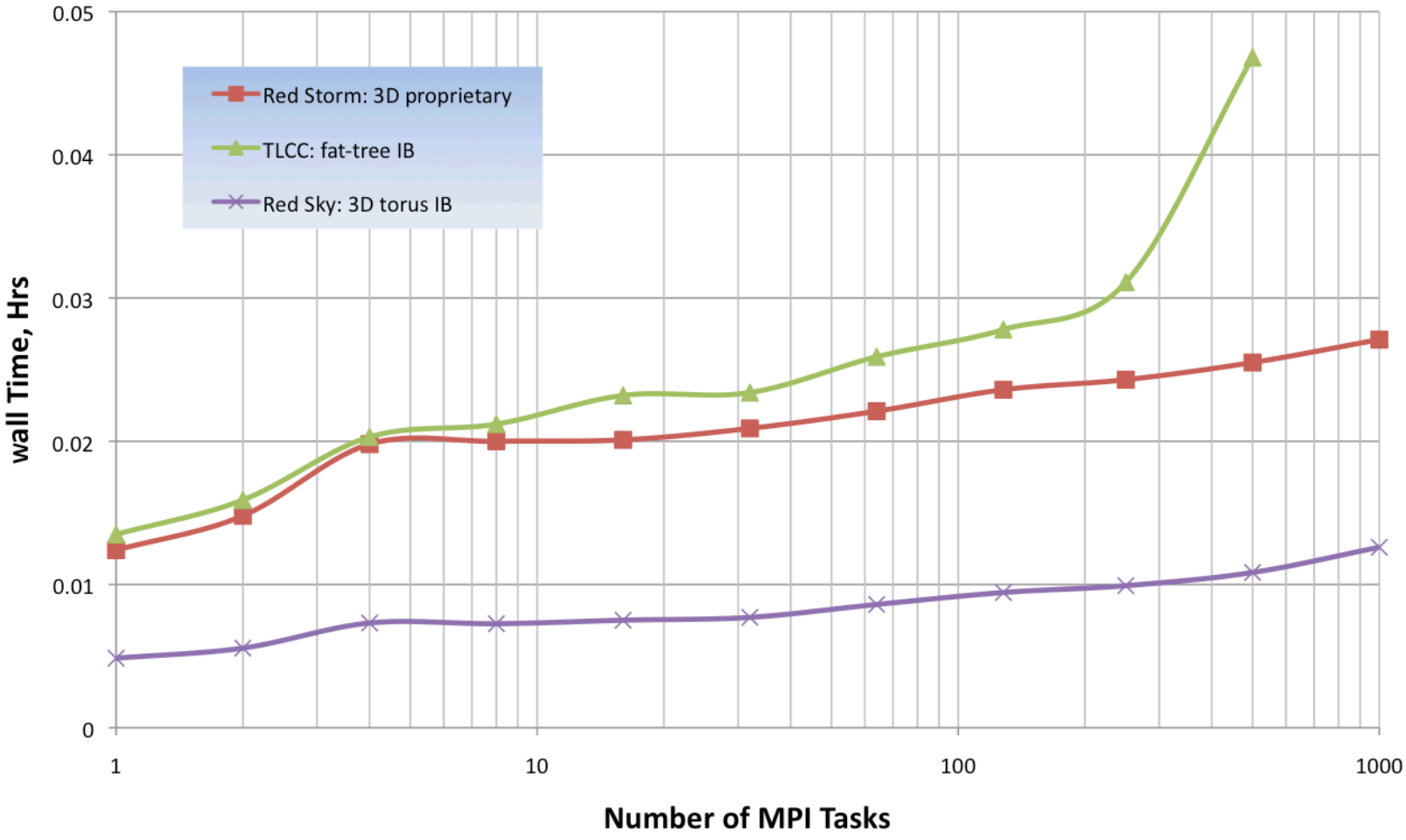
## CTH Shape Charge: Wall Time for 100 time Steps: Weak Scaling with 80x192x80 Cells/core





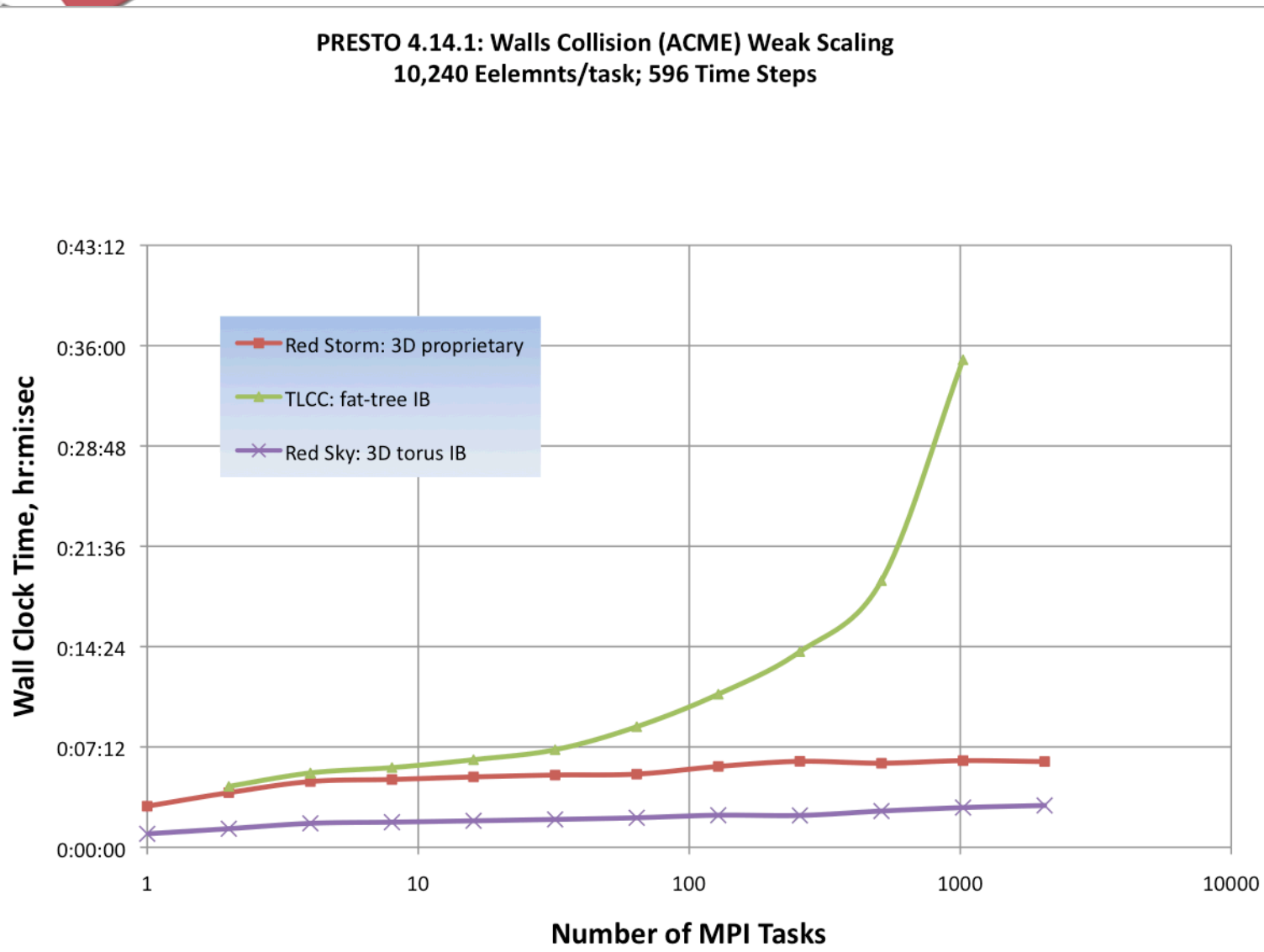

Sandia National Laboratories

# DSMC ICARUS: Fourier Model: Weak Scaling



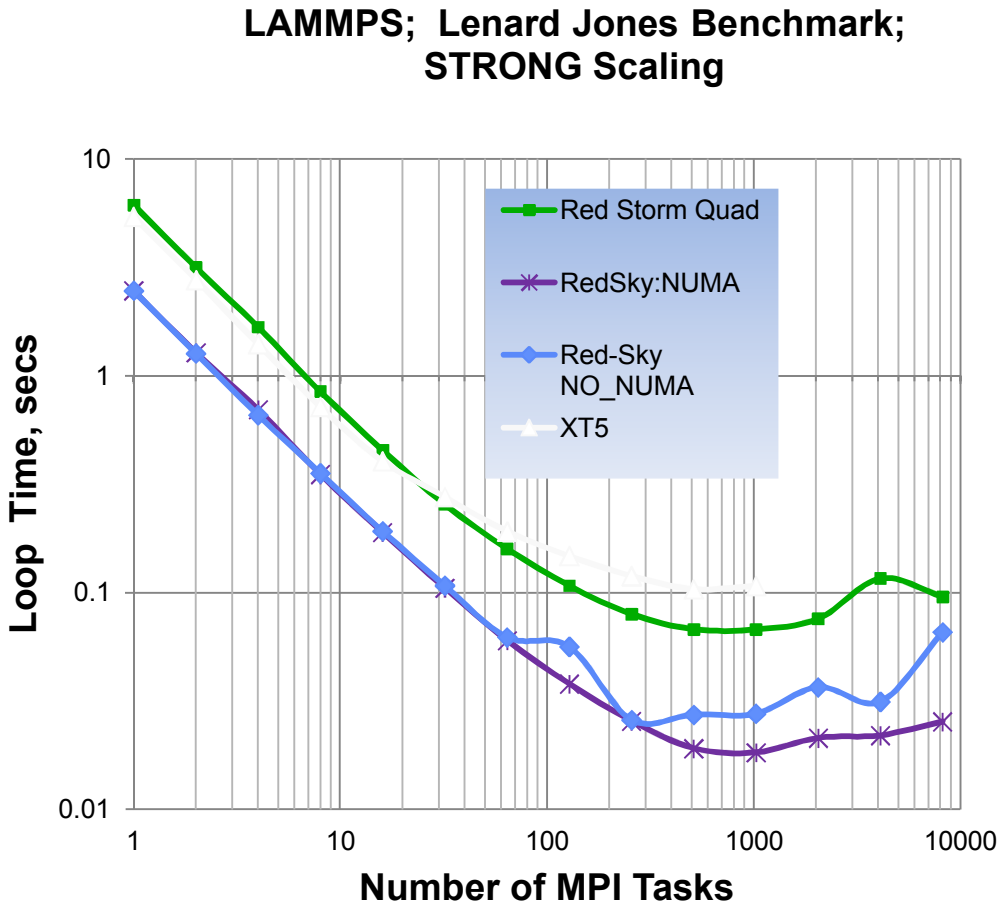



Sandia National Laboratories


PRESTO 4.14.1: Walls Collision (ACME) Weak Scaling
10,240 Eelemnts/task; 596 Time Steps
Wall Clock Time, hr:mi:sec
0:43:12
0:36:00
0:28:48
0:21:36
0:14:24
0:07:12
0:00:00
1
10
100
1000
10000
Number of MPI Tasks
Red Storm: 3D proprietary
TLCC: fat-tree IB
Red Sky: 3D torus IB

# Red sky performance scaling favorably to 10,000 cores

- Classical molecular dynamics
- Model: Lenard Jones benchmark
- Strong scaling analysis with 32,000 atoms for 100 time steps
- LAMMPS divides the computational domain into three dimensional sub-volumes, and makes the sub-volumes as cubic as possible, The amount of data exchanged is proportional to the surface area of the sub-volume.
- Little sensitivity to memory performance
- This example illustrates how even if the application does not stress the memory or the interconnect, Red Storm shows superior scalability






Sandia National Laboratories