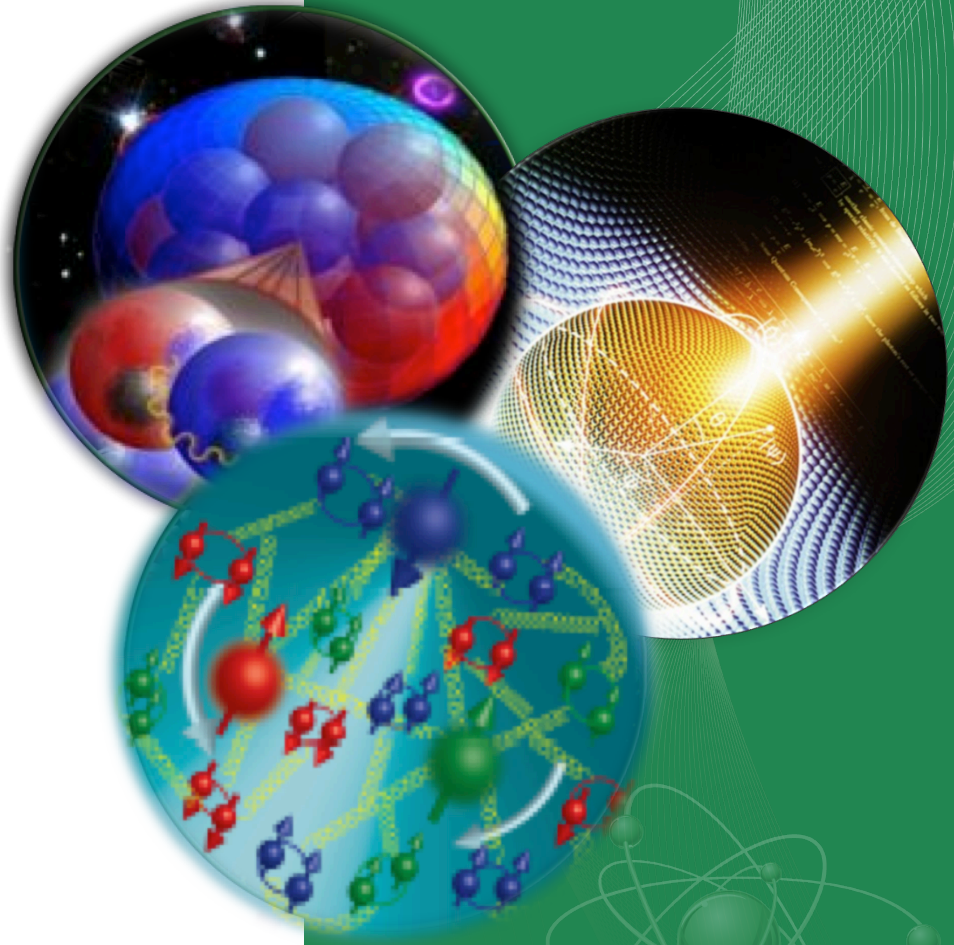
Computational Science



David J. Dean

Director, Physics Division

ORNL

June 24, 2015

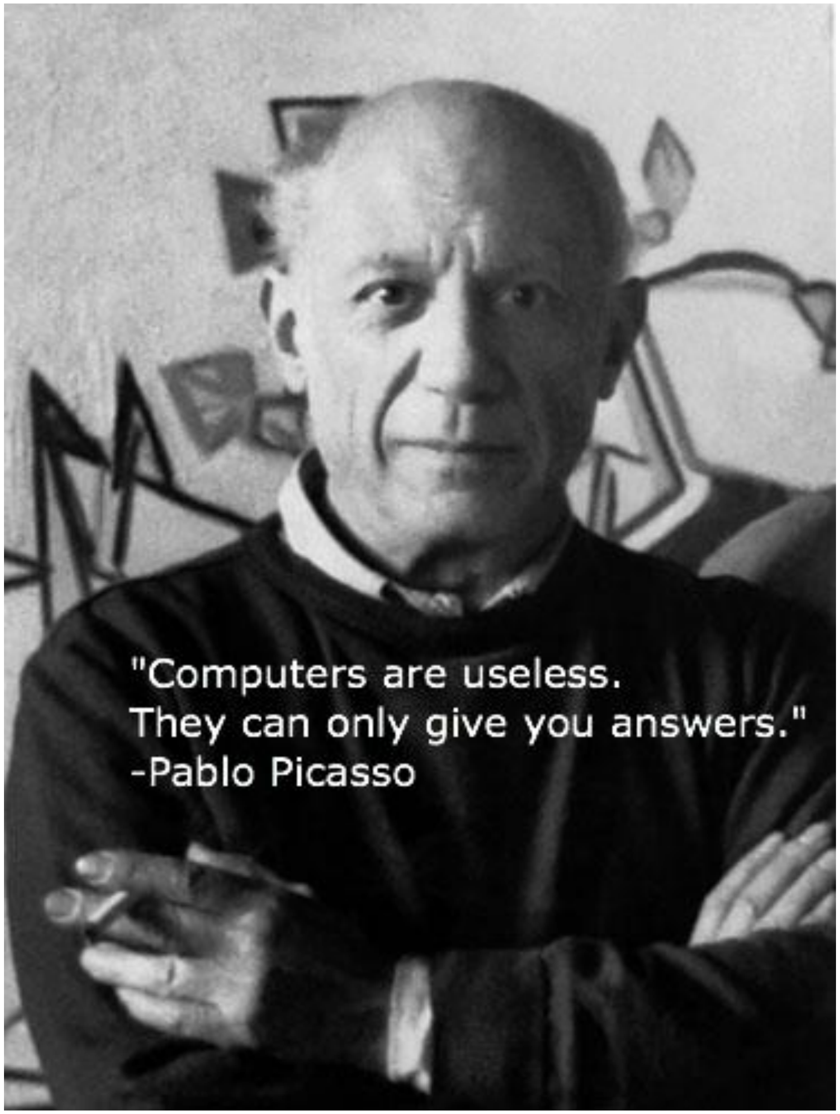
ORNL is managed by UT-Battelle
for the US Department of Energy

Outline

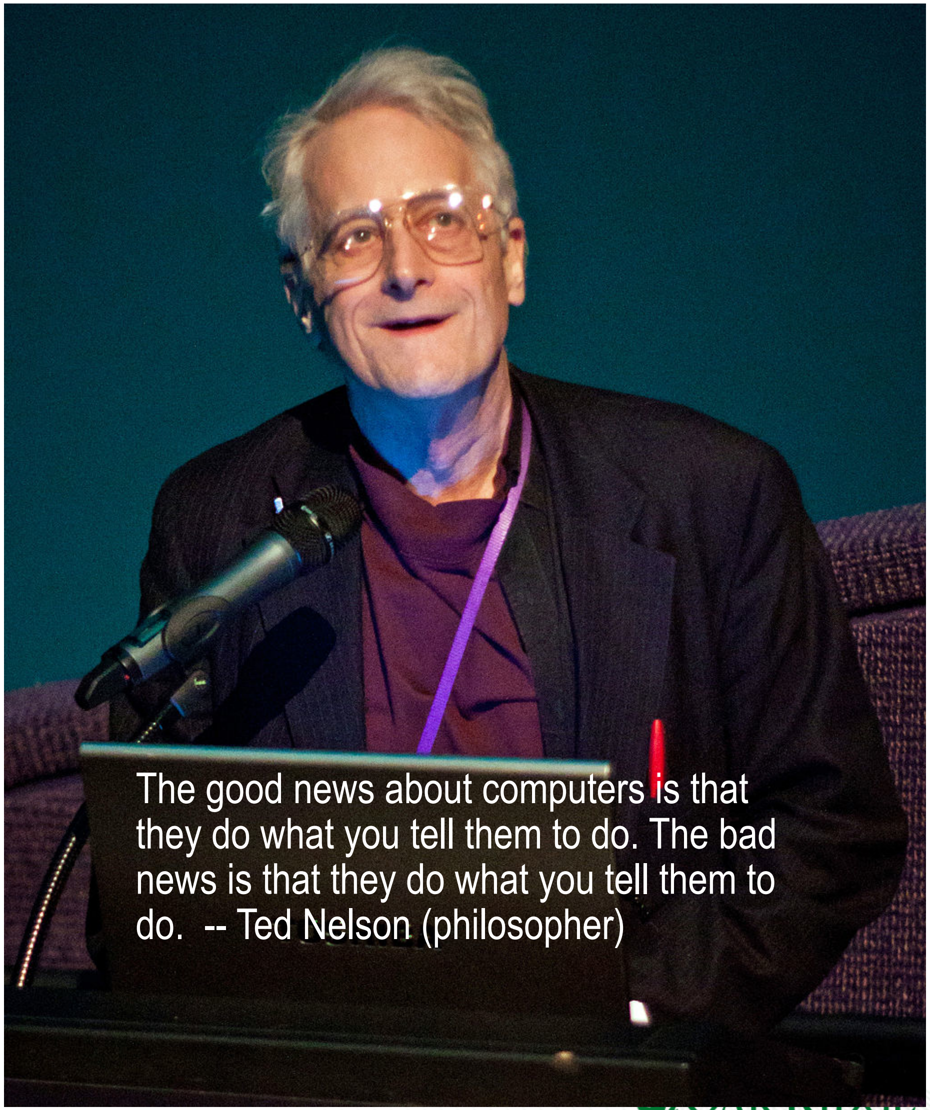
- Why do we compute?
- Some trends in HPC
- Primary computational programs at DOE
- Examples
  - Nuclear Physics
  - Climate and energy R&D
- Beyond exascale

Why do we compute?

Why do we compute?



"Computers are useless.
They can only give you answers."
-Pablo Picasso



The good news about computers is that they do what you tell them to do. The bad news is that they do what you tell them to do. -- Ted Nelson (philosopher)

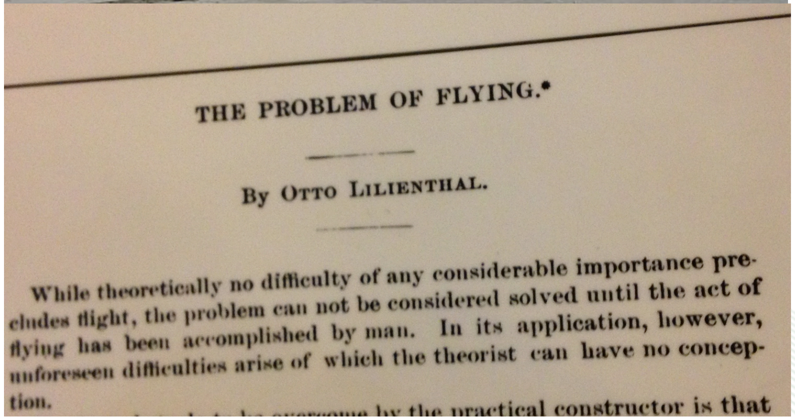
Why and how do we compute?

- Why?
  - Very few instances of analytical, closed form, real life solutions exist.
  - Nonlinearity and emergent behavior exist everywhere.
- We compare theory (as codified in equations) to experiment
- We predict the outcomes of experiments to test theory
- We employ methods of Validation and Verification (V&V)
  - Doing the problem right (numerically sound approaches)
  - Doing the right problem (physically sound approaches)
- We quantify our uncertainties (UQ)
- We apply liberal amounts of physics intuition

A fact of life

Application of the scientific method


The sublime moment of validation

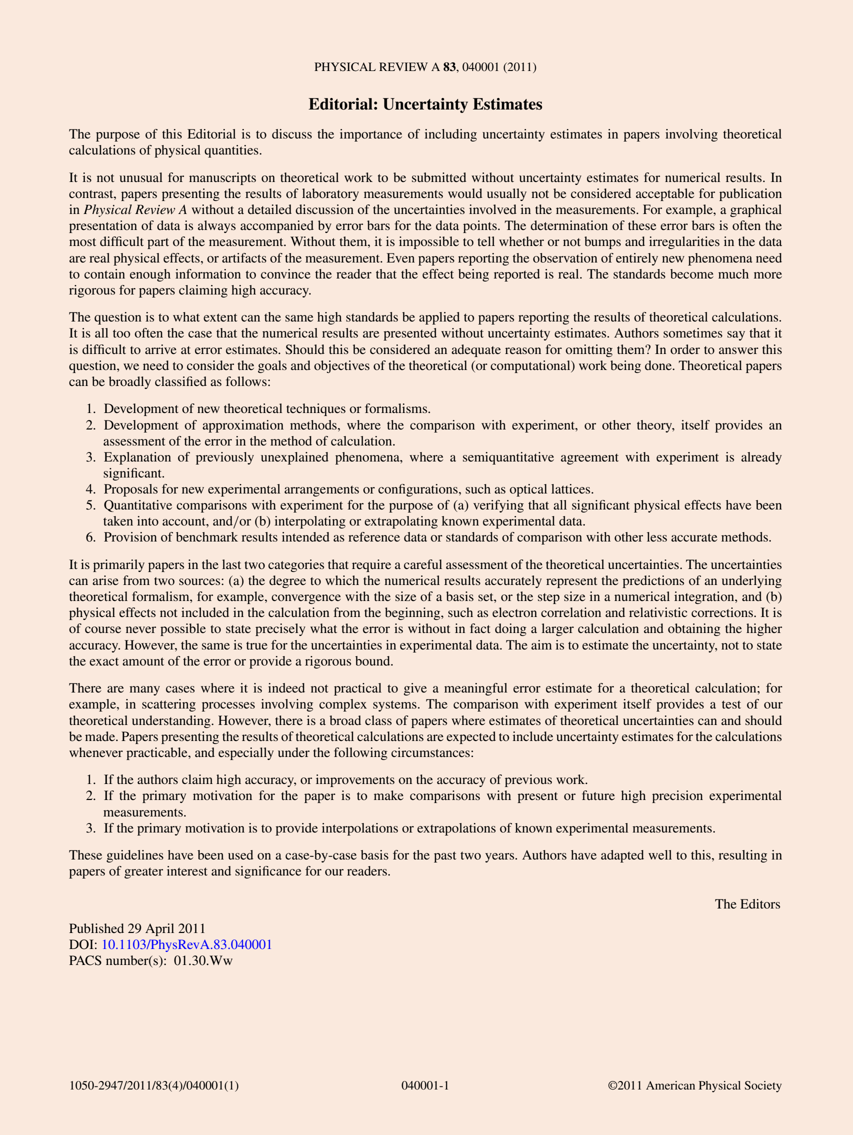

THE PROBLEM OF FLYING.*

By Otto Lilienthal.

While theoretically no difficulty of any considerable importance precludes flight, the problem can not be considered solved until the act of flying has been accomplished by man. In its application, however, unforeseen difficulties arise of which the theorist can have no conception.

by the practical constructor is that

How do you know what you know?



PHYSICAL REVIEW A **83**, 040001 (2011)

Editorial: Uncertainty Estimates

The purpose of this Editorial is to discuss the importance of including uncertainty estimates in papers involving theoretical calculations of physical quantities.

It is not unusual for manuscripts on theoretical work to be submitted without uncertainty estimates for numerical results. In contrast, papers presenting the results of laboratory measurements would usually not be considered acceptable for publication in *Physical Review A* without a detailed discussion of the uncertainties involved in the measurements. For example, a graphical presentation of data is always accompanied by error bars for the data points. The determination of these error bars is often the most difficult part of the measurement. Without them, it is impossible to tell whether or not bumps and irregularities in the data are real physical effects, or artifacts of the measurement. Even papers reporting the observation of entirely new phenomena need to contain enough information to convince the reader that the effect being reported is real. The standards become much more rigorous for papers claiming high accuracy.

The question is to what extent can the same high standards be applied to papers reporting the results of theoretical calculations. It is all too often the case that the numerical results are presented without uncertainty estimates. Authors sometimes say that it is difficult to arrive at error estimates. Should this be considered an adequate reason for omitting them? In order to answer this question, we need to consider the goals and objectives of the theoretical (or computational) work being done. Theoretical papers can be broadly classified as follows:

1. Development of new theoretical techniques or formalisms.
2. Development of approximation methods, where the comparison with experiment, or other theory, itself provides an assessment of the error in the method of calculation.
3. Explanation of previously unexplained phenomena, where a semiquantitative agreement with experiment is already significant.
4. Proposals for new experimental arrangements or configurations, such as optical lattices.
5. Quantitative comparisons with experiment for the purpose of (a) verifying that all significant physical effects have been taken into account, and/or (b) interpolating or extrapolating known experimental data.
6. Provision of benchmark results intended as reference data or standards of comparison with other less accurate methods.

It is primarily papers in the last two categories that require a careful assessment of the theoretical uncertainties. The uncertainties can arise from two sources: (a) the degree to which the numerical results accurately represent the predictions of an underlying theoretical formalism, for example, convergence with the size of a basis set, or the step size in a numerical integration, and (b) physical effects not included in the calculation from the beginning, such as electron correlation and relativistic corrections. It is of course never possible to state precisely what the error is without in fact doing a larger calculation and obtaining the higher accuracy. However, the same is true for the uncertainties in experimental data. The aim is to estimate the uncertainty, not to state the exact amount of the error or provide a rigorous bound.

There are many cases where it is indeed not practical to give a meaningful error estimate for a theoretical calculation; for example, in scattering processes involving complex systems. The comparison with experiment itself provides a test of our theoretical understanding. However, there is a broad class of papers where estimates of theoretical uncertainties can and should be made. Papers presenting the results of theoretical calculations are expected to include uncertainty estimates for the calculations whenever practicable, and especially under the following circumstances:

1. If the authors claim high accuracy, or improvements on the accuracy of previous work.
2. If the primary motivation for the paper is to make comparisons with present or future high precision experimental measurements.
3. If the primary motivation is to provide interpolations or extrapolations of known experimental measurements.

These guidelines have been used on a case-by-case basis for the past two years. Authors have adapted well to this, resulting in papers of greater interest and significance for our readers.

The Editors

Published 29 April 2011
DOI: 10.1103/PhysRevA.83.040001
PACS number(s): 01.30.Ww

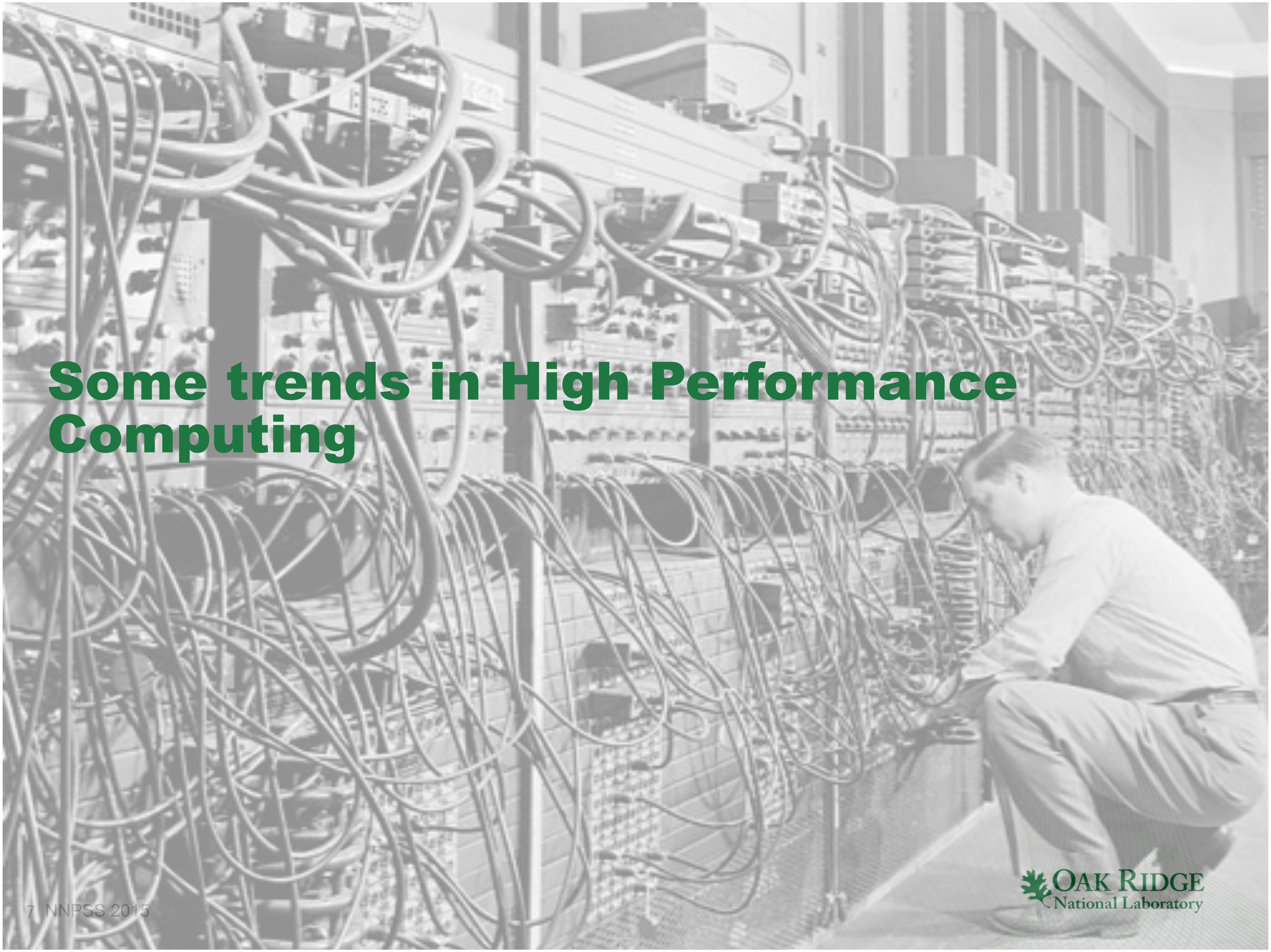
1050-2947/2011/83(4)/040001(1) 040001-1

'It is not unusual for manuscripts on theoretical work to be submitted without uncertainty estimates for numerical results'

'Papers presenting the results of theoretical calculations are expected to include uncertainty estimates for the calculations whenever practicable.'

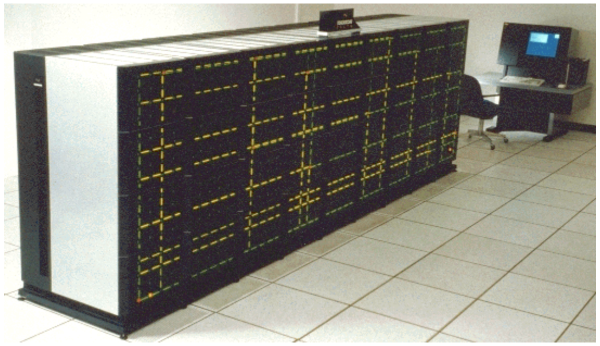
- Claim of high accuracy
- Comparison with high precision experimental measurements
- Interpolation or extrapolation of known experimental measurements

Phys. Rev. A 83, 040001 (2011)
(atomic, molecular, optical physics)



Some trends in High Performance Computing

Some open science computers...

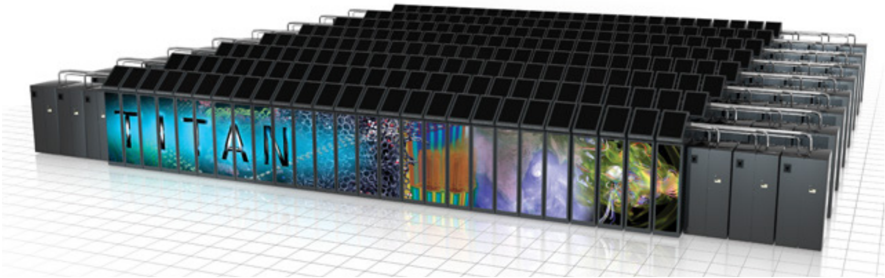


Touchstone Delta – Caltech
June 1993
#8 on Top500
Rpeak=13.9 Gflop/s (#1 = 59.7 GF)

NERSC T3E900
June 1998
#8
321 Gflop/s (#1 = 1,338 GF)

NERSC, IBM SP-3, 16 way
June 2003
#4
7 Tflop/s (#1 = 35.86 TF)

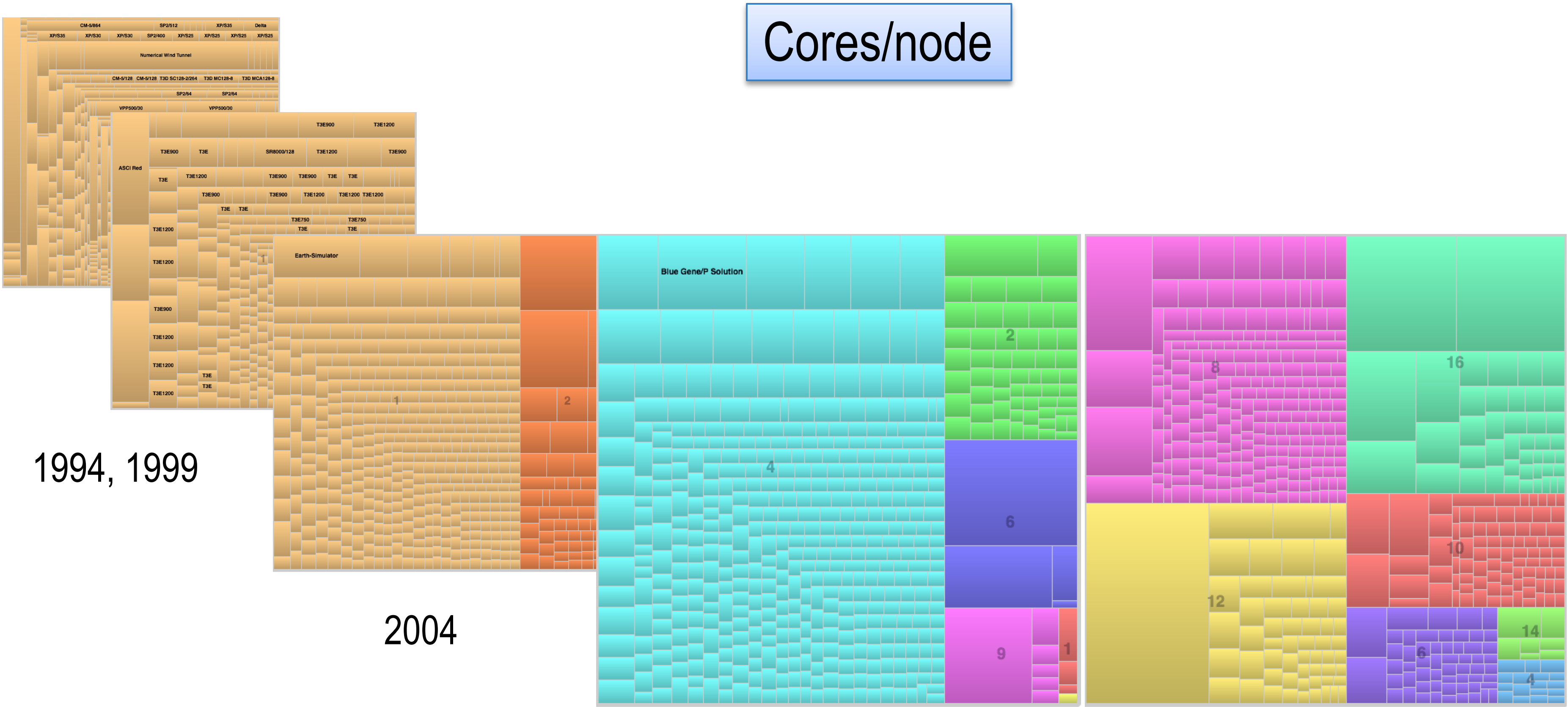
ORNL, Cray XT4 (upgraded components)
June 2008
#6
205 Tflop/s (#1 = 1.026 PF)



ORNL, Cray XK7
November 2014
#2
17.5 Pflop/s (#1 = 33.9 Pflop/s)

Changes across 21 years…

Cores/node



1994, 1999

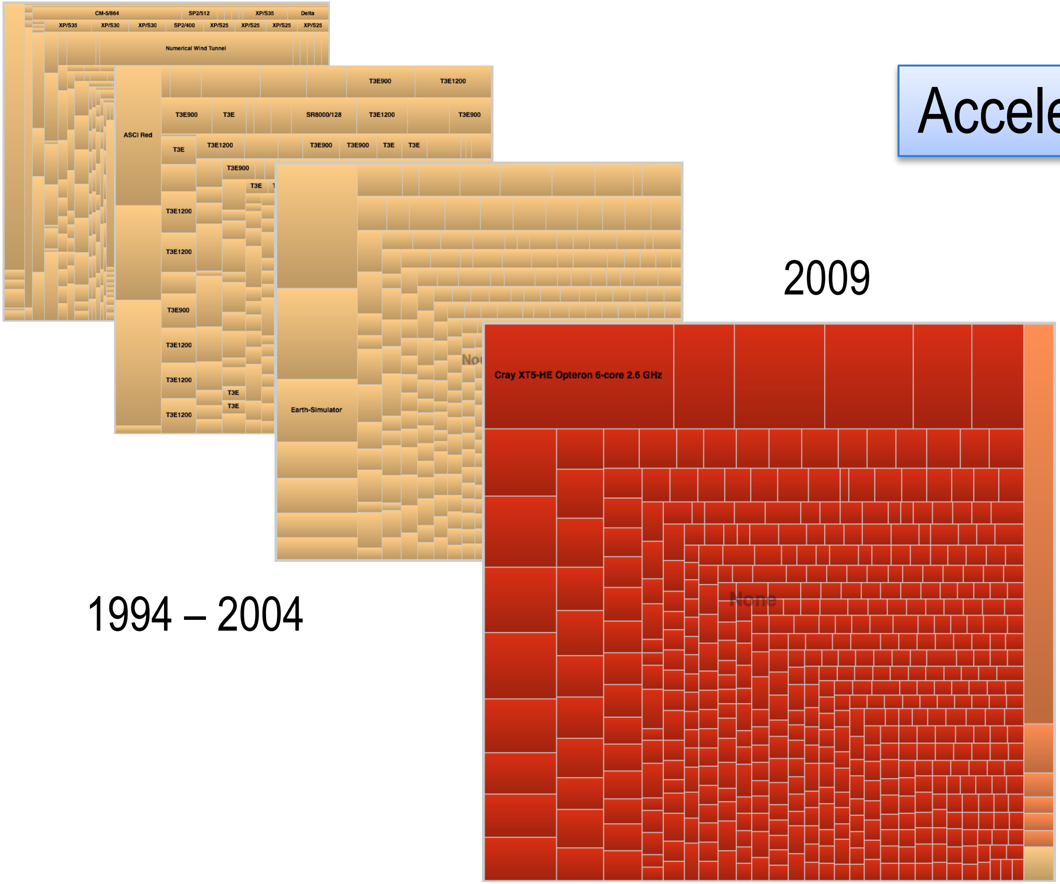
2004

20009

2014

MPI/OpenMP/multi threading…

Changes across 21 years…

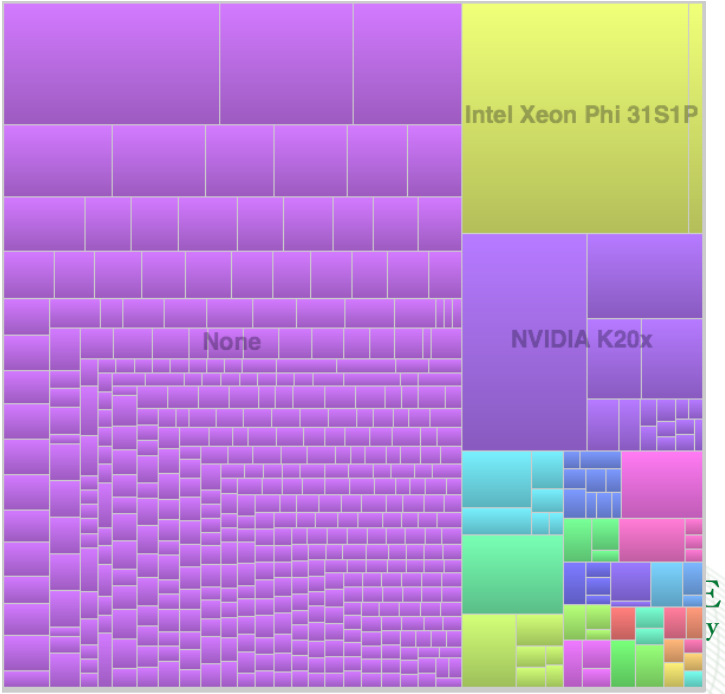


1994 – 2004

2009

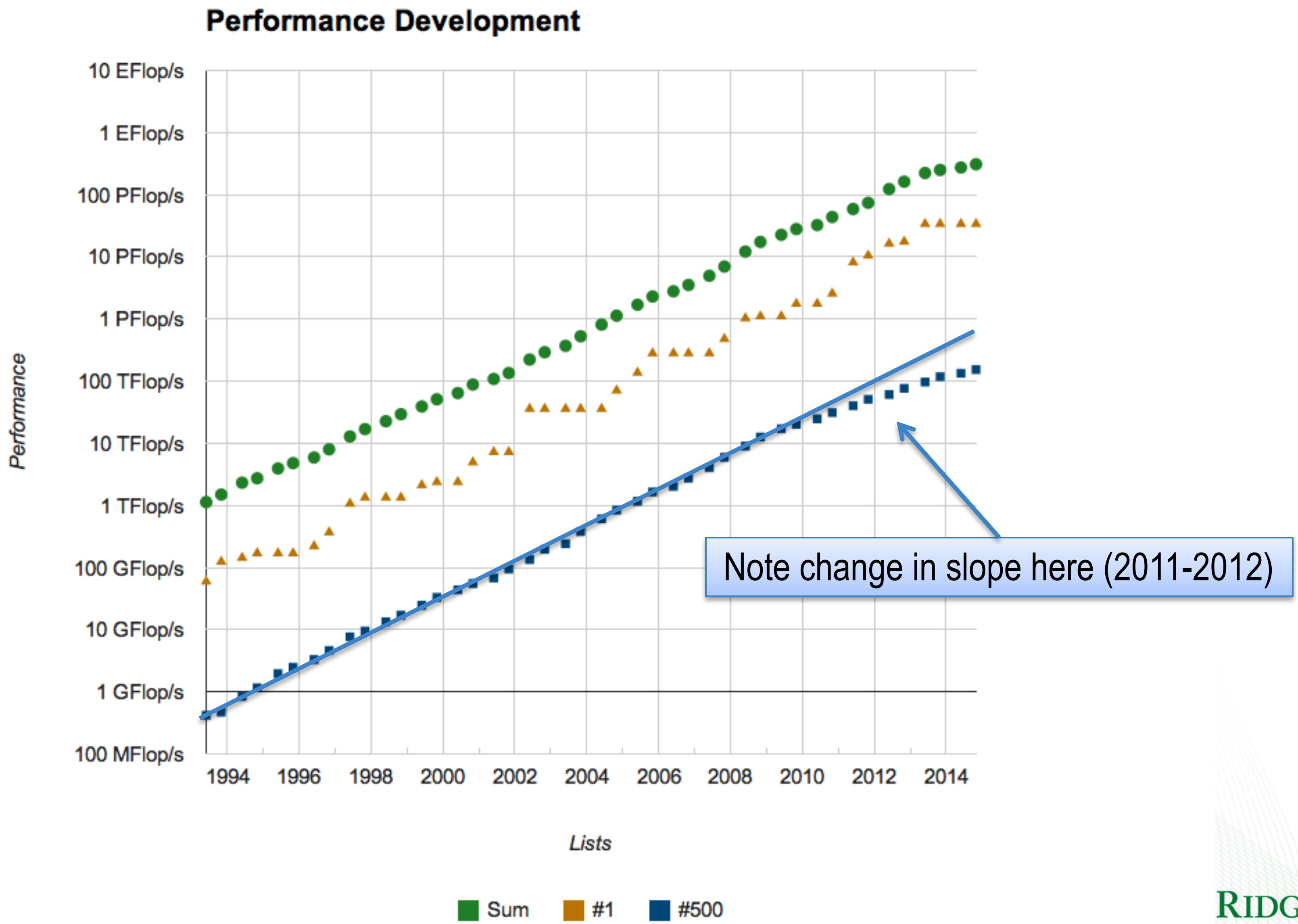
Accelerators: CPU/GPU

2014



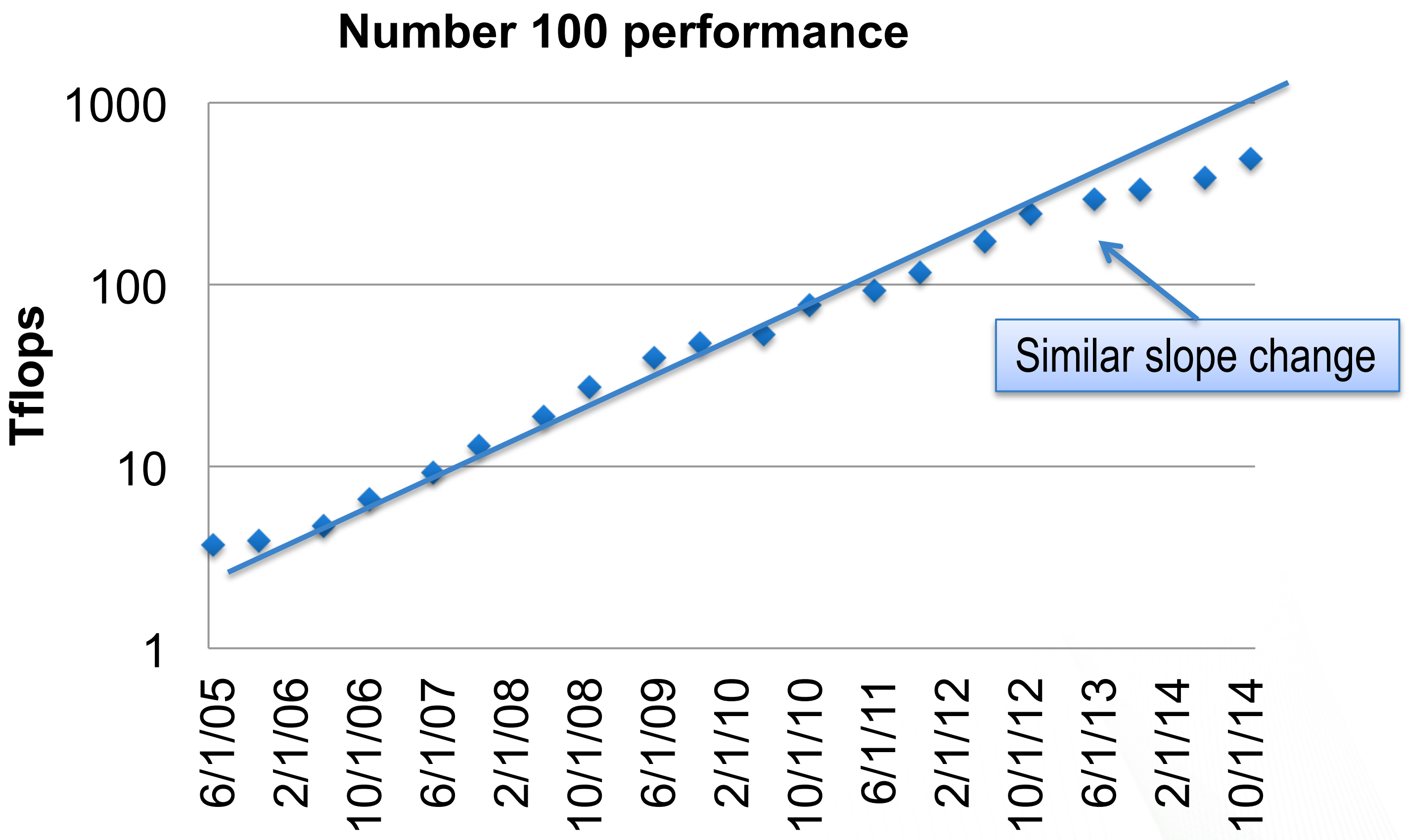
MPI/OpenMP/multi threading/CUDA…

Development with time



Note change in slope here (2011-2012)

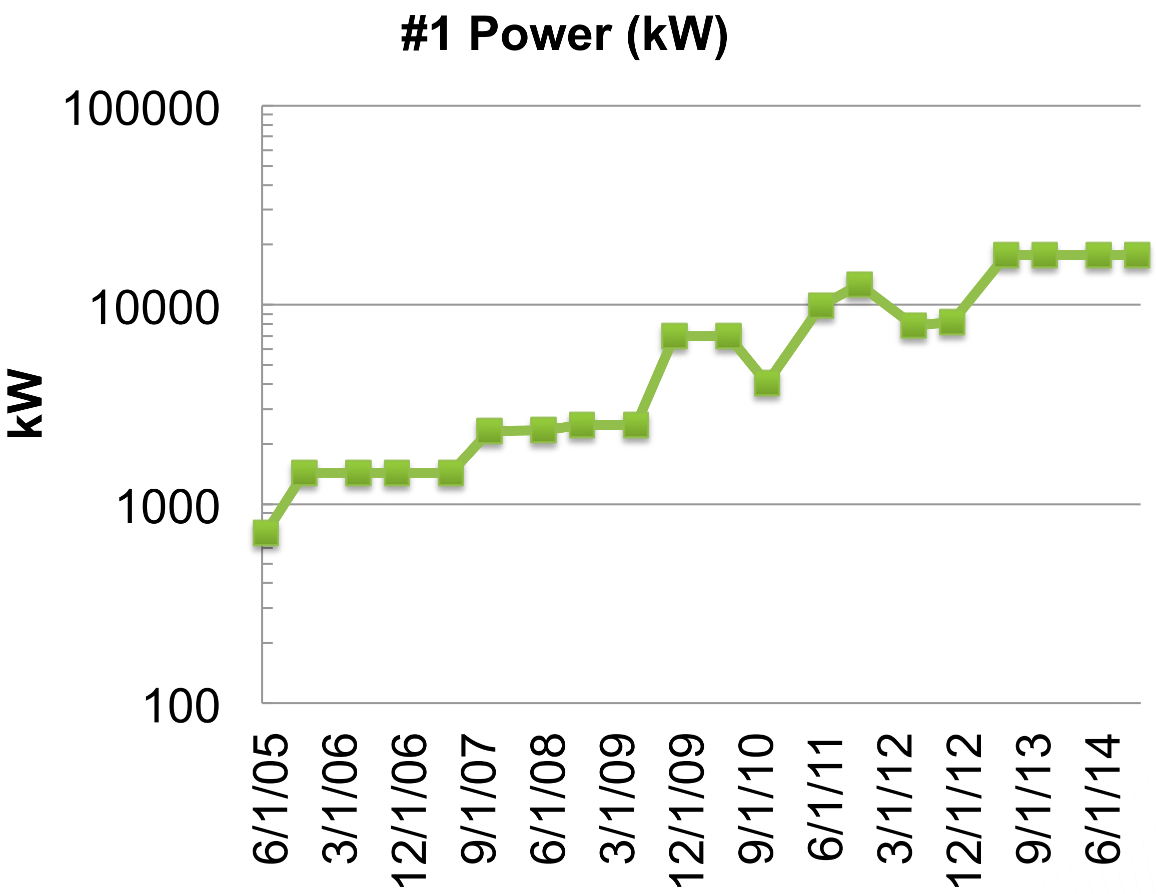
Performance of #100



Similar slope change

Slope is changing…happens later for higher performing systems…
Implies a longer doubling time for most systems

A big issue: power



Incremental cost of running RHIC: $550k/week

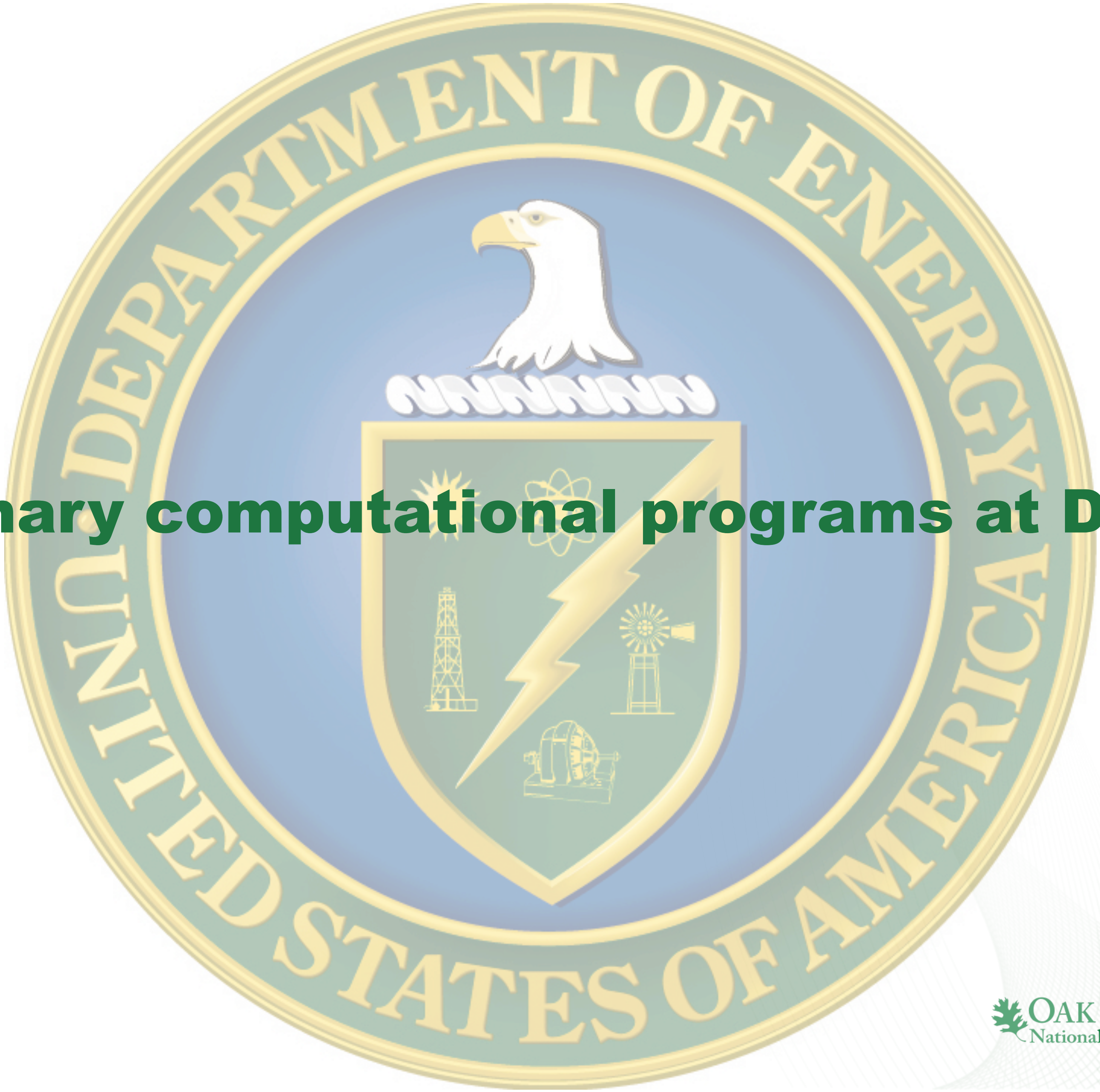
Incremental cost of running Titan: $140k/week

Incremental cost of running Tianhe-2: $300k/week

(assume $0.1/kW-h)

June 2005 Tflop/kW = 0.191
Nov. 2014 Tflop/kW = 1.901

10x technology improvement



Primary computational programs at DOE

Advanced Simulations and Computing

Part of the Science Based Stockpile Stewardship Program at NNSA

- Ensuring nuclear weapon reliability, safety, and performance in the absence of testing
- Program incorporates:
    - Integrated codes
    - Physics and engineering models
    - Computational systems and software environments
    - Facility operations (at LLNL, LANL and SNL)

ASCI Red, 1 TF

ASCI White, 7.2 TF

BG/L, 70.7 TF

Road runner, 1.1 PF

Sequoia, 16.3 PF

Advanced Scientific Computing Research

ASCR
Advanced Scientific Computing Research

U.S. DEPARTMENT OF ENERGY | Office of Science

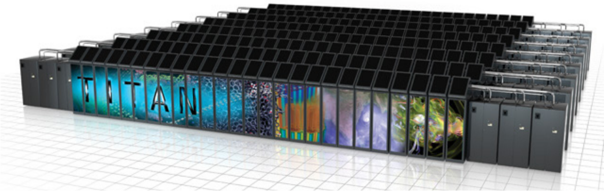
Mission: discover, develop, and deploy computational and networking capabilities to analyze, model, simulate and predict complex phenomena important to the DOE

- Applied Math & Computer Science
- Facilities operations

SciDAC
Scientific Discovery through Advanced Computing

NERSC

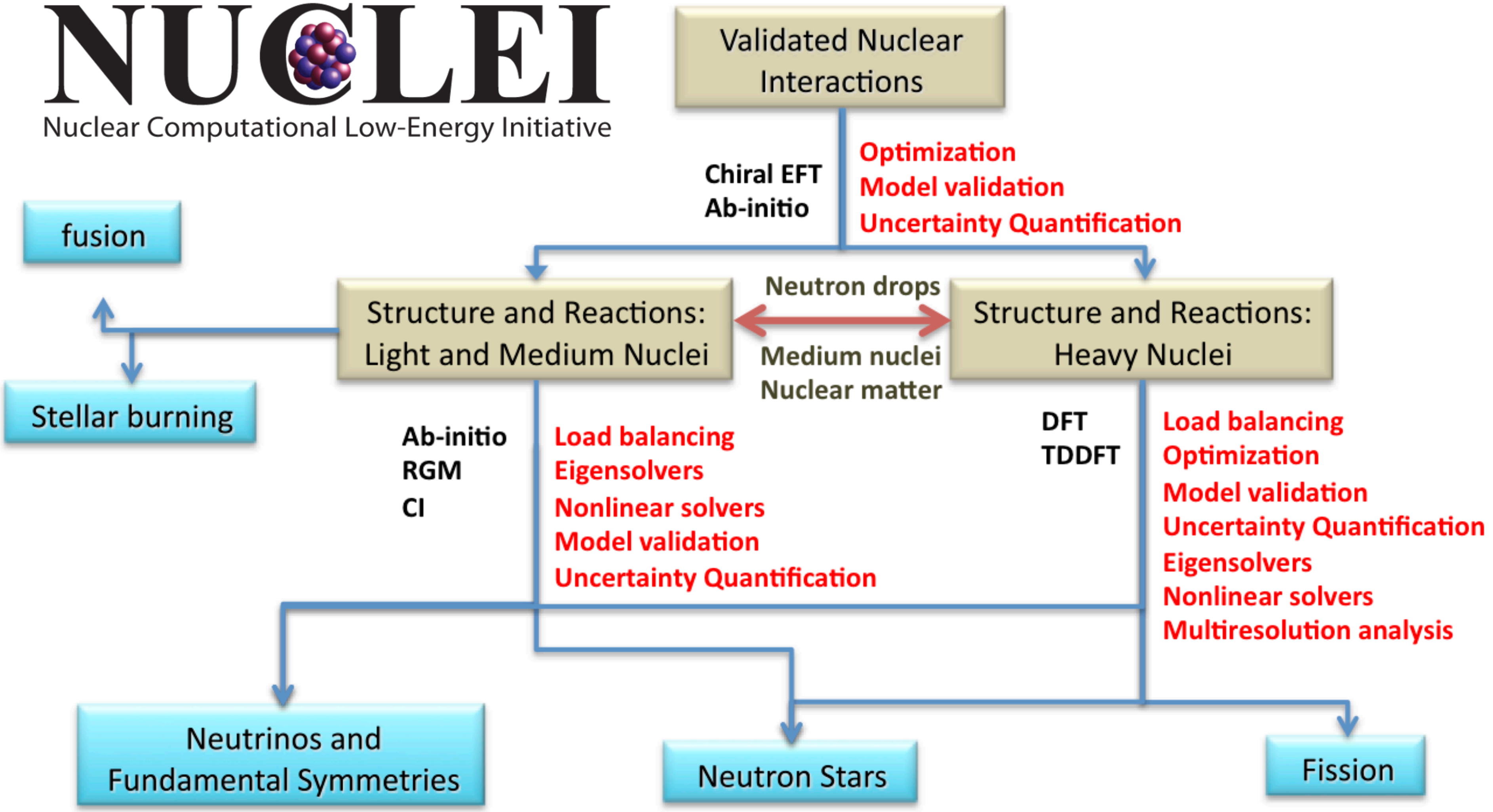
OLCF
OAK RIDGE LEADERSHIP COMPUTING FACILITY



Argonne NATIONAL LABORATORY | Leadership Computing Facility

SciDAC example

NUCLEI
Nuclear Computational Low-Energy Initiative



Validated Nuclear Interactions

Chiral EFT
Ab-initio

Optimization
Model validation
Uncertainty Quantification

Structure and Reactions: Light and Medium Nuclei

Neutron drops
Medium nuclei
Nuclear matter

Structure and Reactions: Heavy Nuclei

fusion

Stellar burning

Ab-initio
RGM
CI

Load balancing
Eigensolvers
Nonlinear solvers
Model validation
Uncertainty Quantification

DFT
TDDFT

Load balancing
Optimization
Model validation
Uncertainty Quantification
Eigensolvers
Nonlinear solvers
Multiresolution analysis

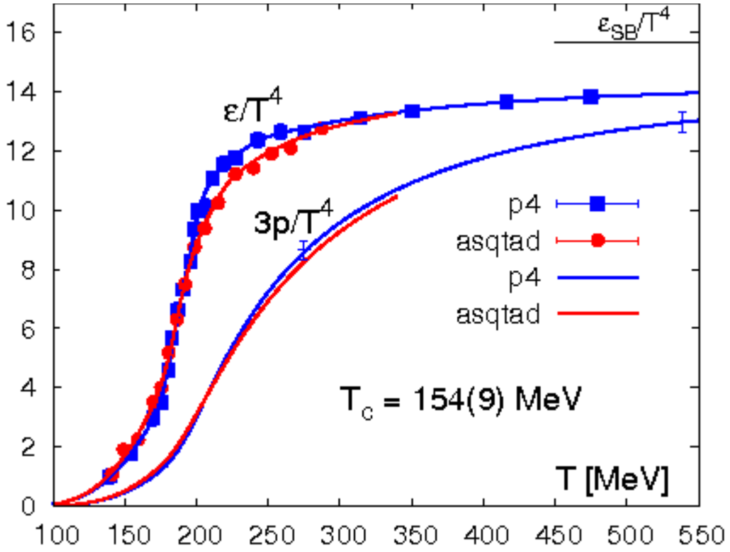
Neutrinos and Fundamental Symmetries

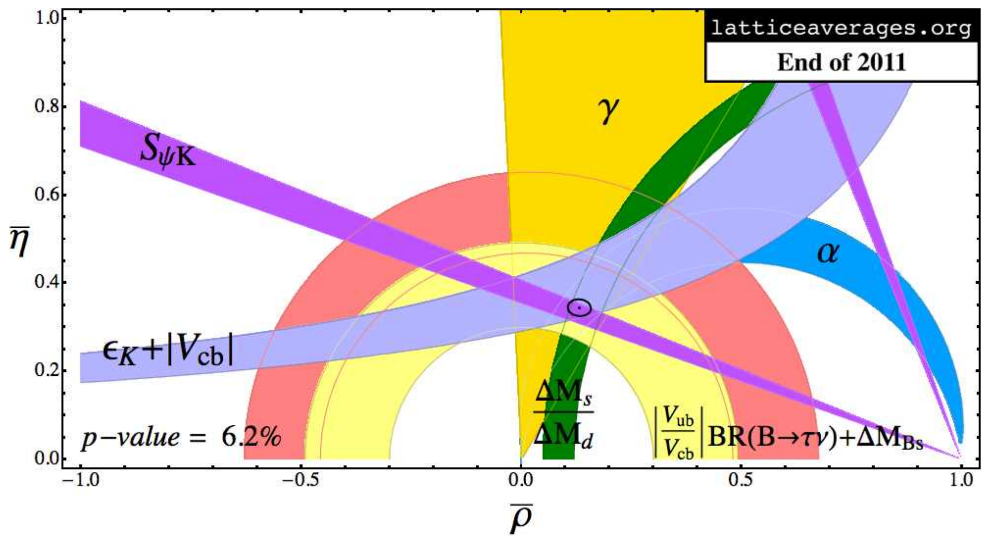
Neutron Stars

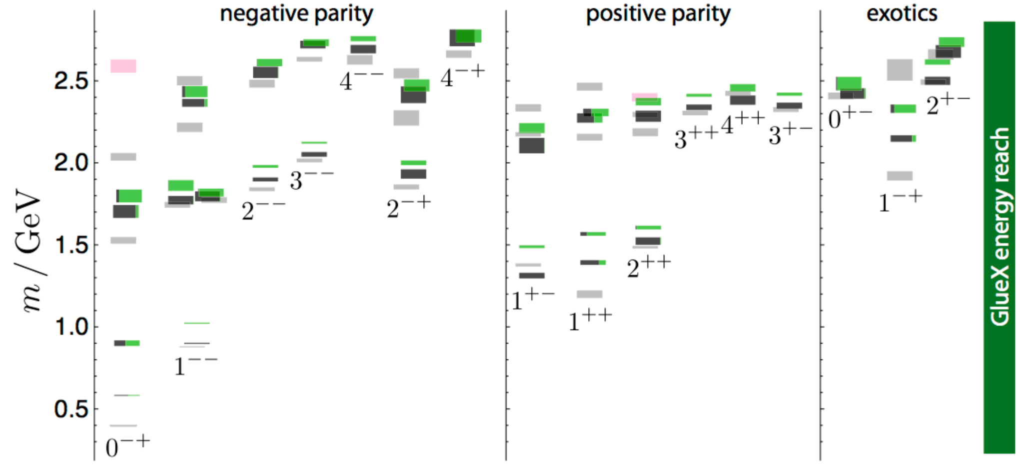
Fission

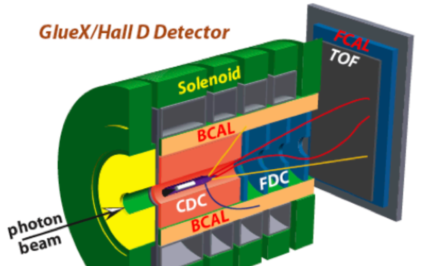
SciDAC example

Computing properties of hadrons, nuclei and nuclear matter from QCD



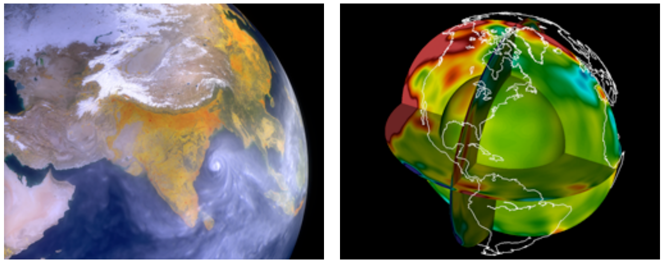




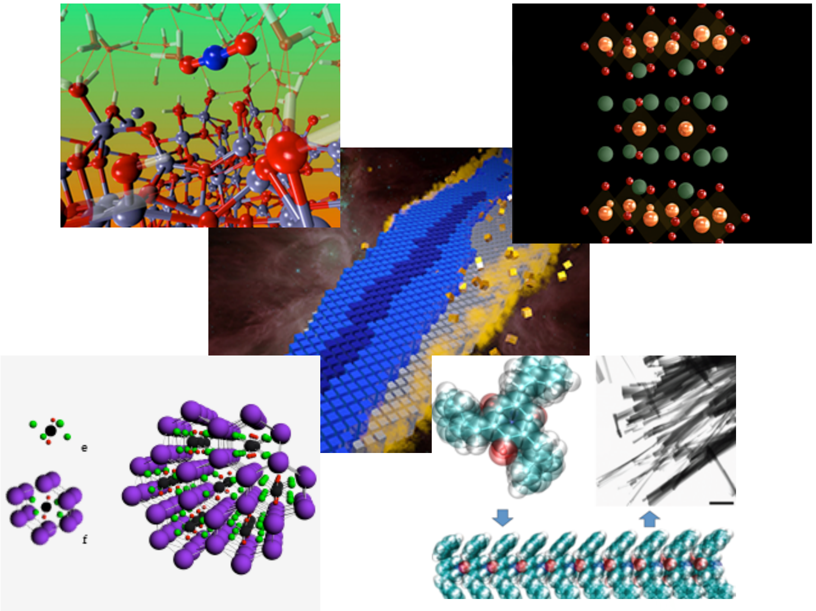


Science at 100-200 Pflop:

Center for Accelerated Application Readiness (CAAR) program

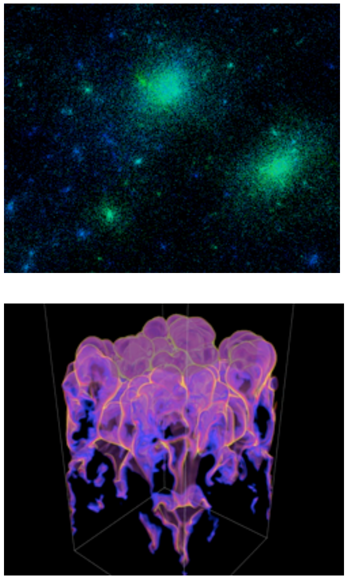


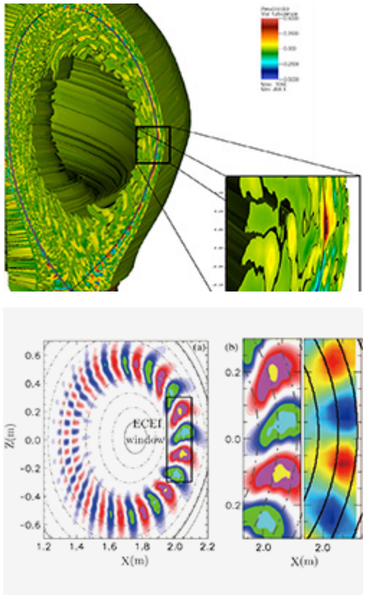
Climate and seismology



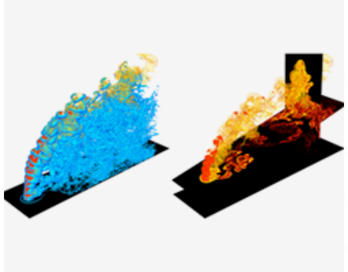
Quantum many body

Astro



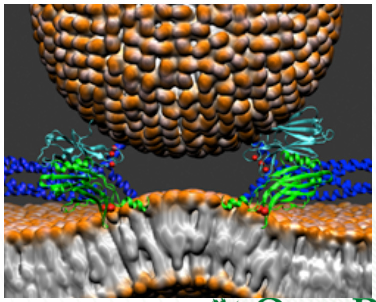


Fusion



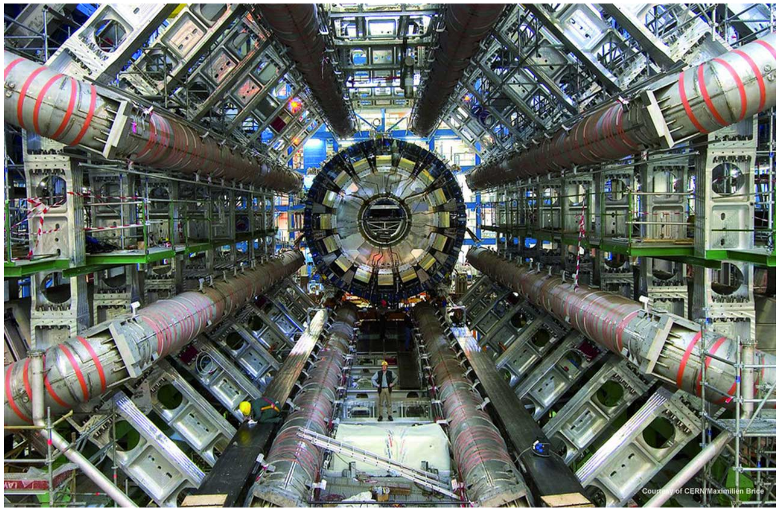
Combustion

Brain chemistry



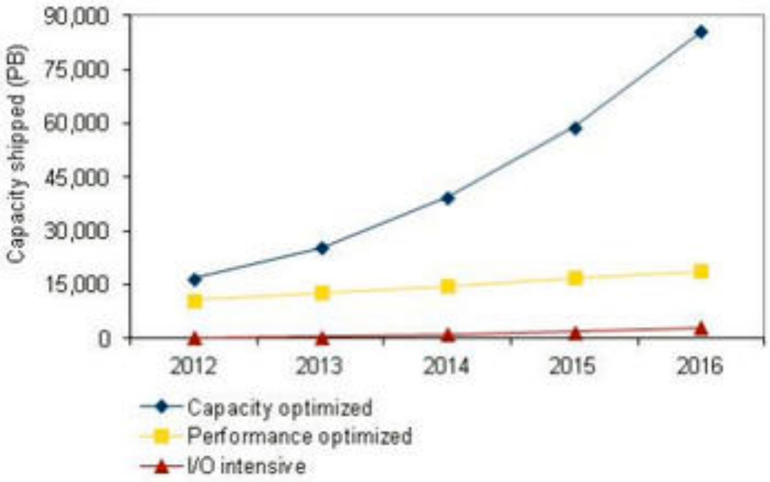
Data and computing

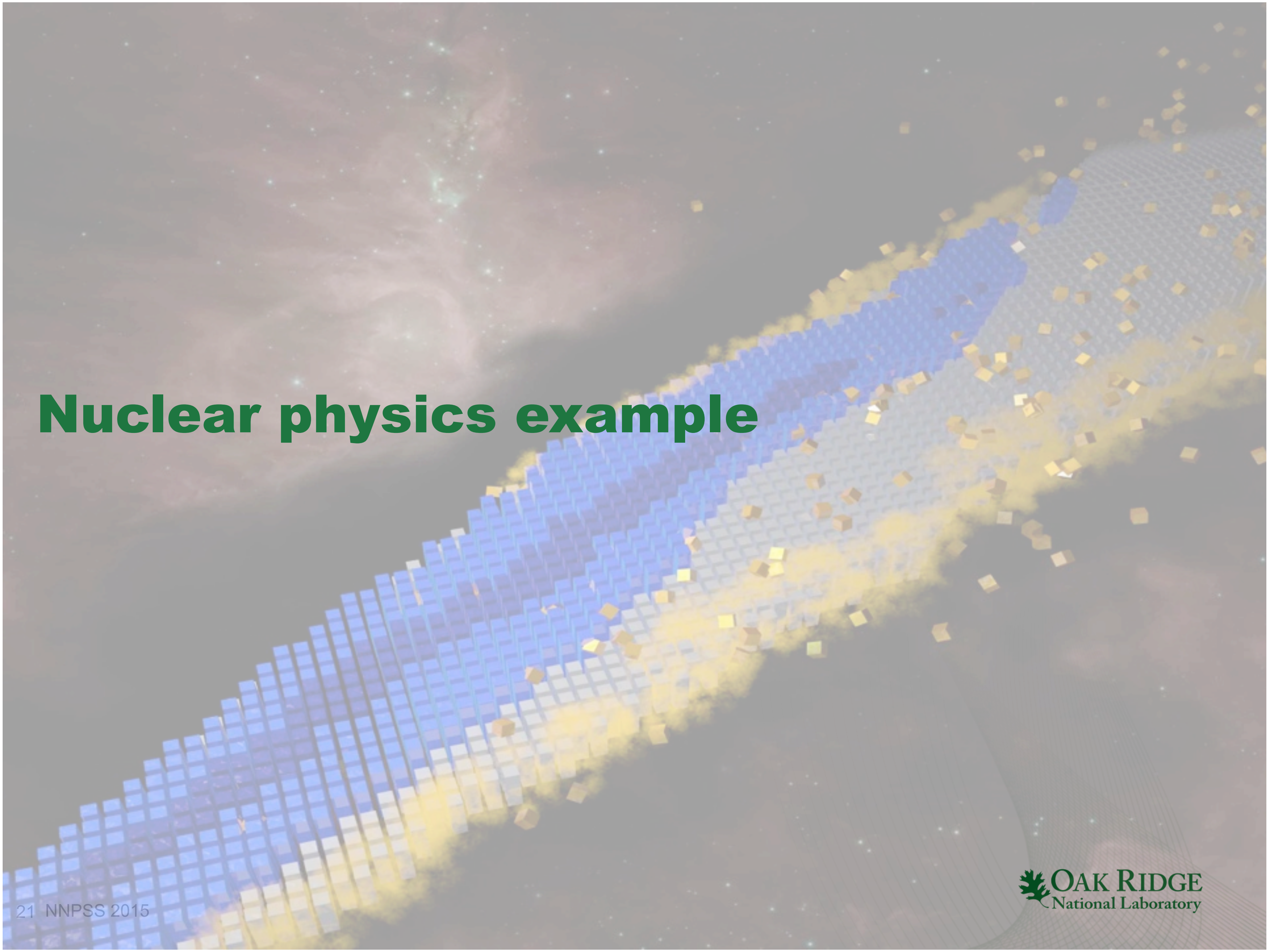
- Characteristics
  - Volume Example: 100x data from LHC after luminosity and energy upgrades
  - Variety: From LHC to 0nubb
  - Velocity: Flow of data from both experiments and simulations
  - Variability: signal to noise
  - Veracity: how good is the data?
- Data presents significant research opportunities



Courtesy of CERN/Maximilien Brice

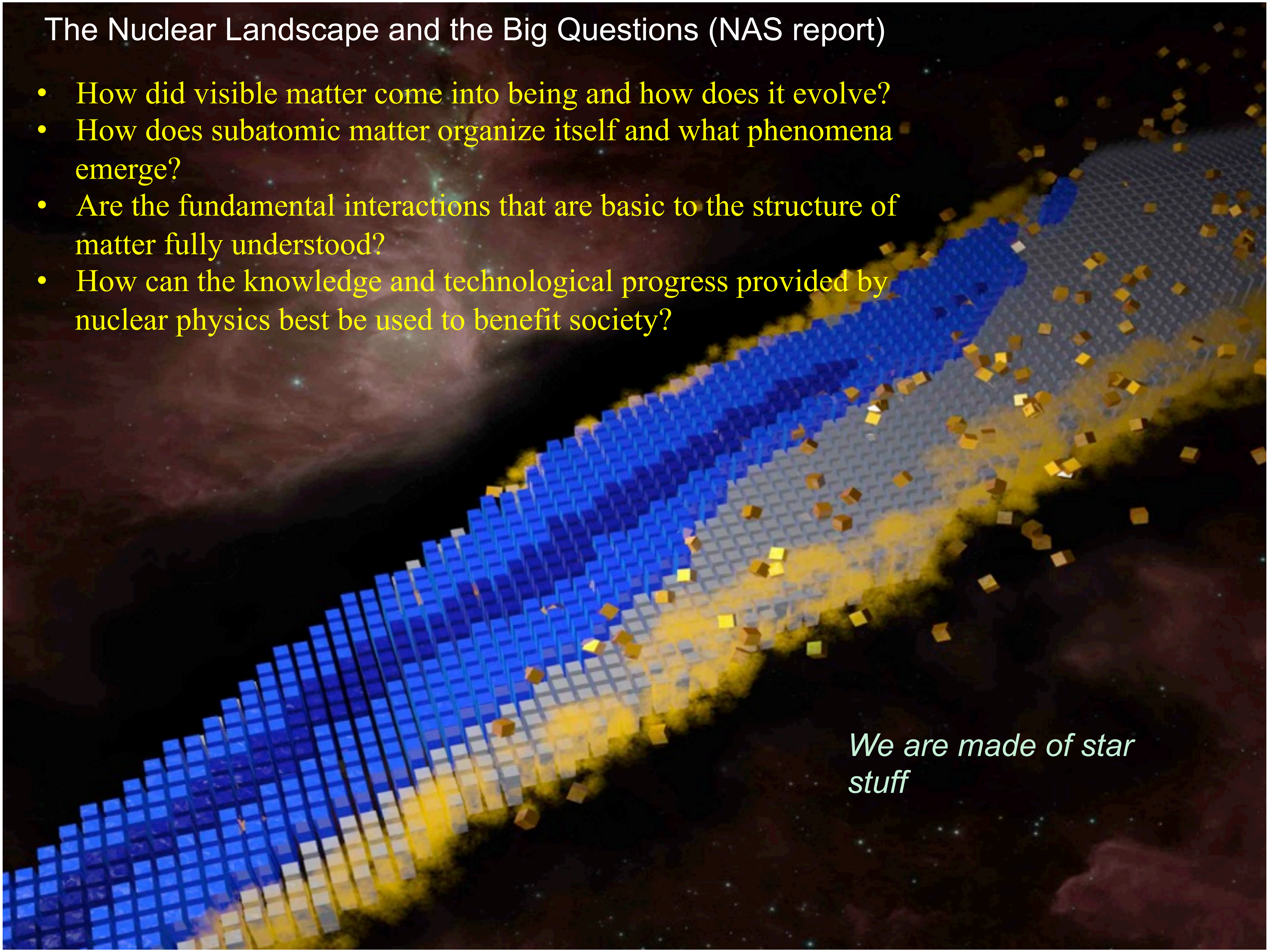
Worldwide Capacity-Optimized, Performance-Optimized, and I/O-Intensive Storage Systems Shipments, 2012–2016





Nuclear physics example

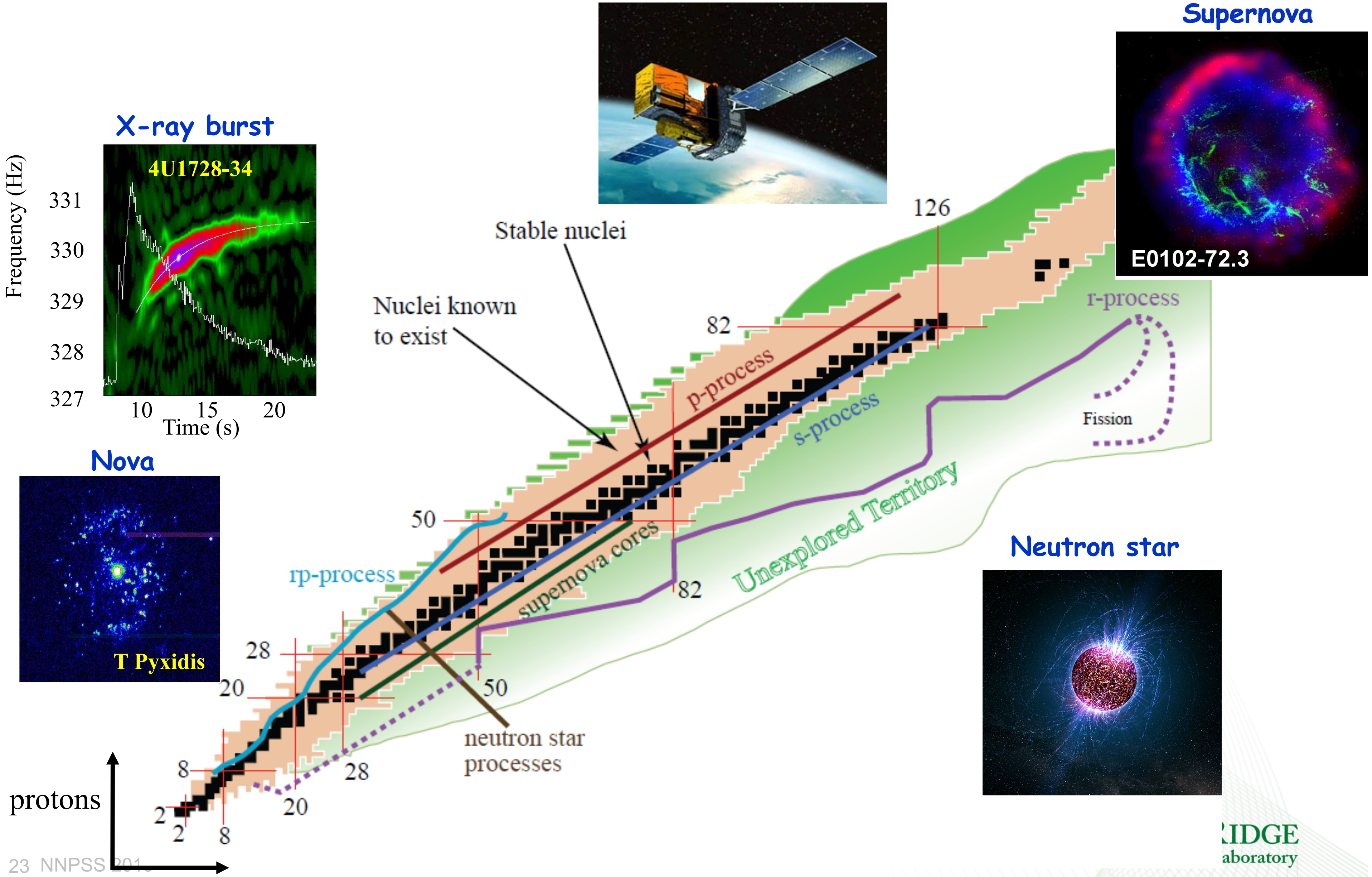
The Nuclear Landscape and the Big Questions (NAS report)



- How did visible matter come into being and how does it evolve?
- How does subatomic matter organize itself and what phenomena emerge?
- Are the fundamental interactions that are basic to the structure of matter fully understood?
- How can the knowledge and technological progress provided by nuclear physics best be used to benefit society?

We are made of star stuff

The nuclear landscape in the cosmos



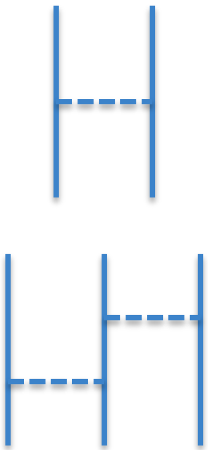
Approaching weakly bound nuclei with coupled cluster theory

A method that captures the physics

$$|\Psi\rangle = \exp(T)|\Phi\rangle$$

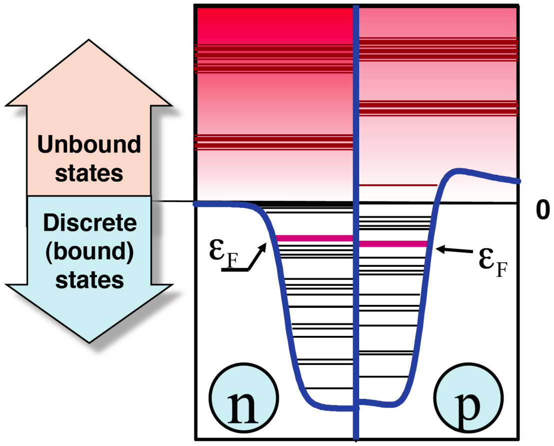
- Coupled cluster theory
- Infinitely summed lower class (1, 2, 3 loop) many-body perturbation theory diagrams
- Amenable to HPC applications

Effective Field Theory for nuclear force (interactions)



- Effective field theory expansion of the nucleon forces that respects symmetries of QCD
- 2-body and 3-body forces

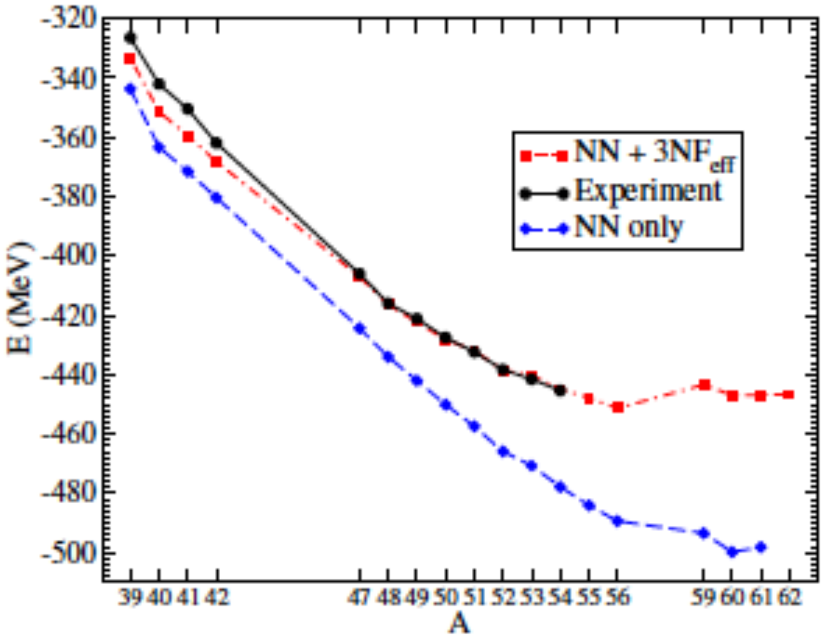
Basis states that incorporate continuum effects

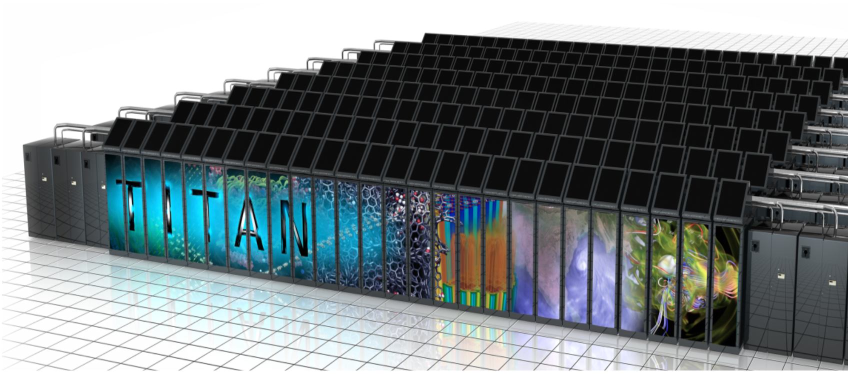


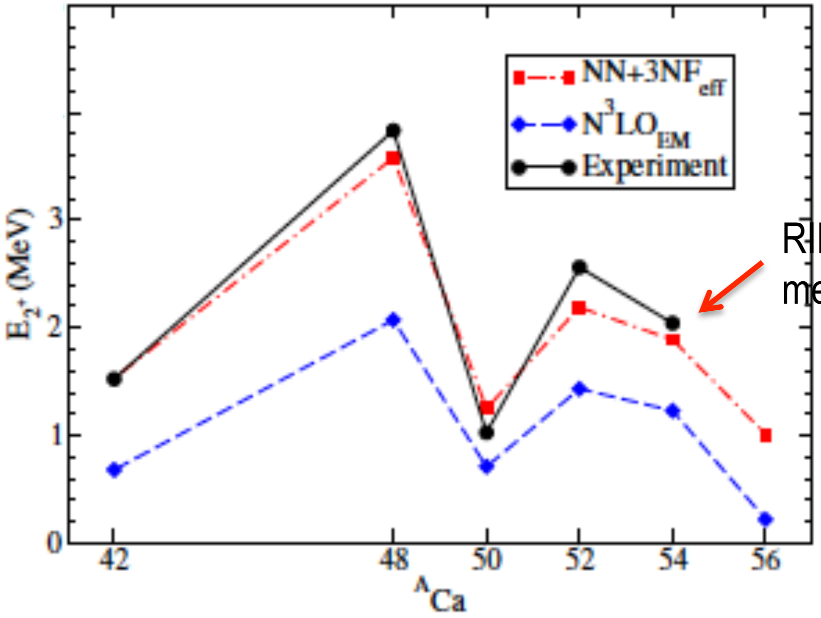
- Basis includes bound, scattering, and continuum states
- Berggren basis

Dean & Hjorth-Jensen, PRC69, 054320 (2004); Kowalski et al., PRL 92, 132501 (2004); Wloch et al., PRL94, 212501 (2005) Gour et al., PRC (2006); Hagen et al, PLB (2006); PRC 2007a, 2007b; Dean, Phys. Today (Nov, 2007)

Investigating weakly bound nuclei







RIKEN measurement

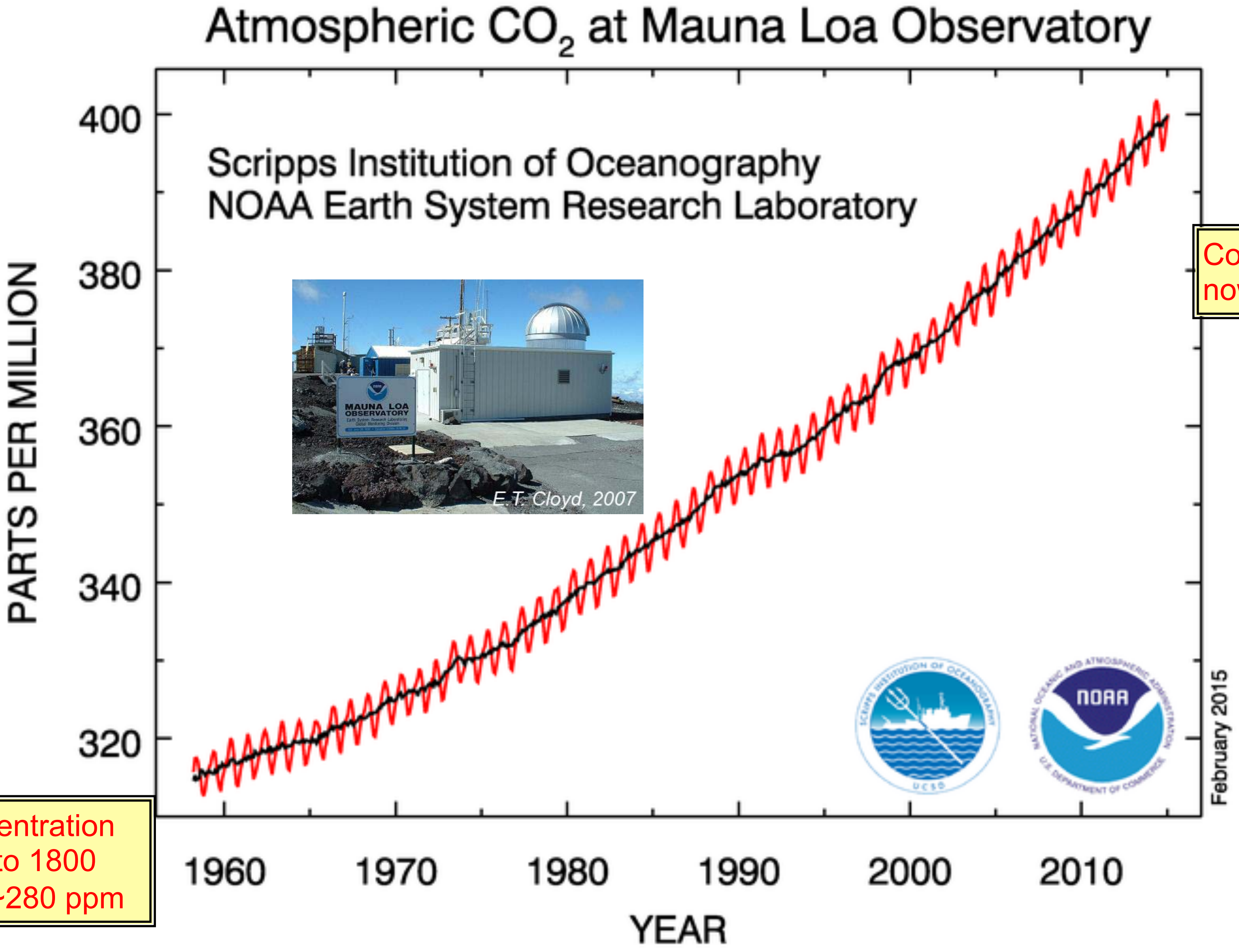
How does one approach the drip line (quickly, smoothly, or asymptotically)?

Are there new shell structures beyond our standard nuclear magic numbers?

Rep. Prog. Phys. 77, 096302 (2014)

The environment, energy and R&D

Observed CO_2 concentrations today



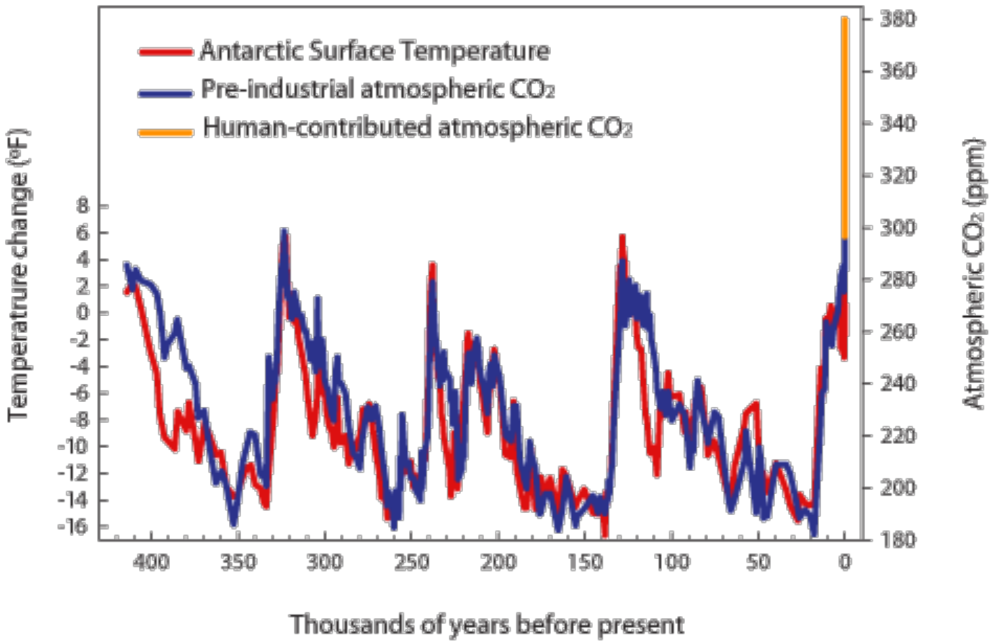
Concentration now ~400 ppm

Concentration prior to 1800 was ~280 ppm

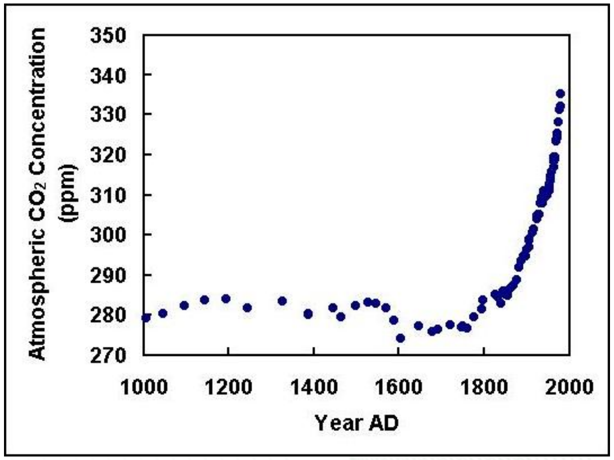
Paleoclimatology

Trends in Atmospheric CO_2 & Global Surface Temperature

The last 400,000 Years



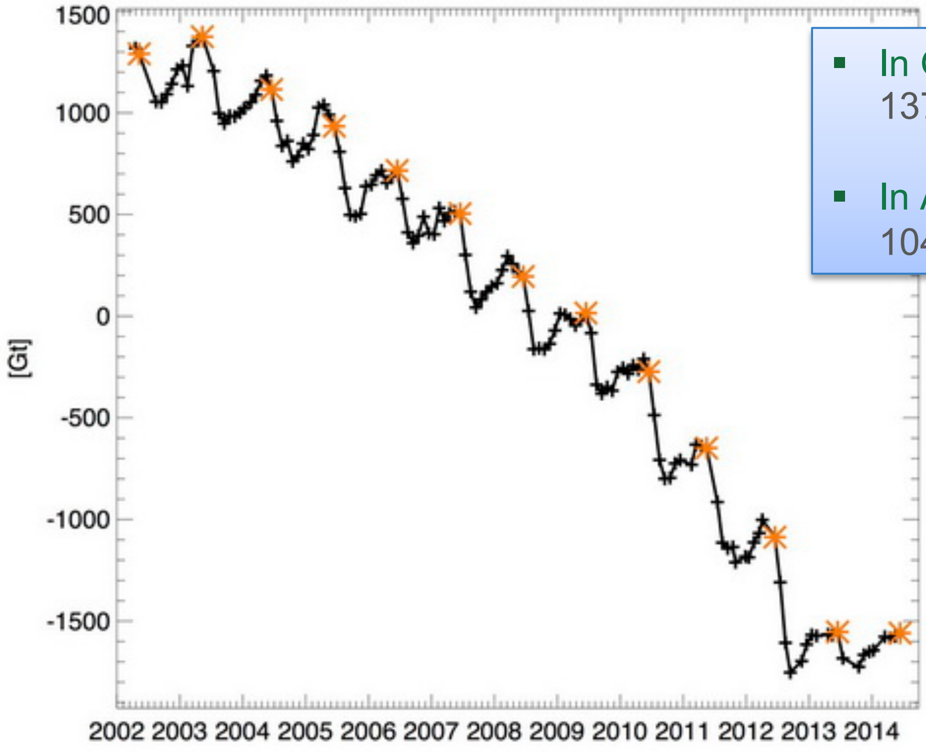
Data Sources:
Atmospheric CO_2 prior to 3000 years ago and Antarctic Surface temperature prior to 100 years ago: J.R., Petit, Jouzel J., et al. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature 399:429-436.
Pre-industrial CO_2 40-3000 years ago: Indermühle A., T.F. Stocker, F., et. al. 1999, Holocene carbon-cycle dynamics based on CO_2 trapped in ice at Taylor Dome, Antarctica. Nature 398, 121-126.
Modern CO_2: Keeling, C.D. and T.P. Whorf. 2005. Atmospheric CO_2 records from sites in the SIO air sampling network. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.



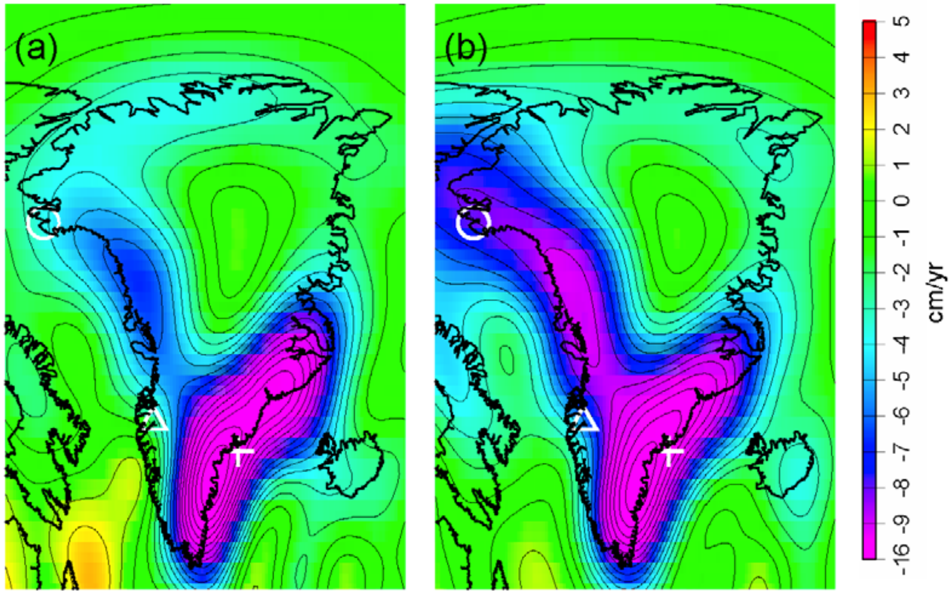
Observed Greenland Ice Loss 2002-2014

Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE (Gravity Recovery and Climate Experiment) satellite:

- In Greenland, the mass loss increased from 137 Gt/yr in 2002–2003 to 286 Gt/yr in 2007–2009
- In Antarctica, the mass loss increased from 104 Gt/yr in 2002–2006 to 246 Gt/yr in 2006–2009

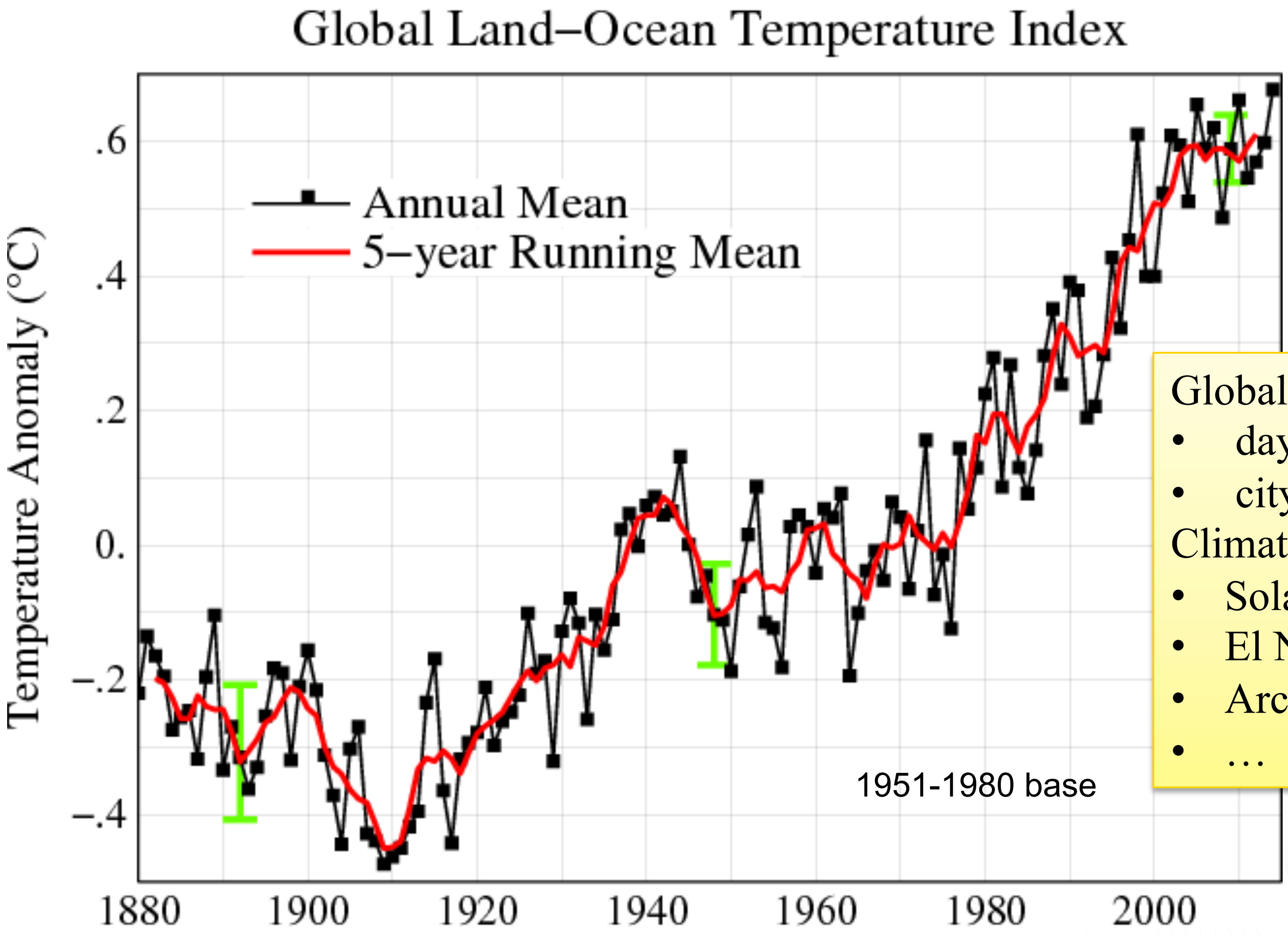


Monthly mass anomalies (in Gigatonnes) for the Greenland ice sheet since April 2002 estimated from GRACE measurements



http://www.arctic.noaa.gov/reportcard/greenland_ice_sheet.html

Observed global temperature



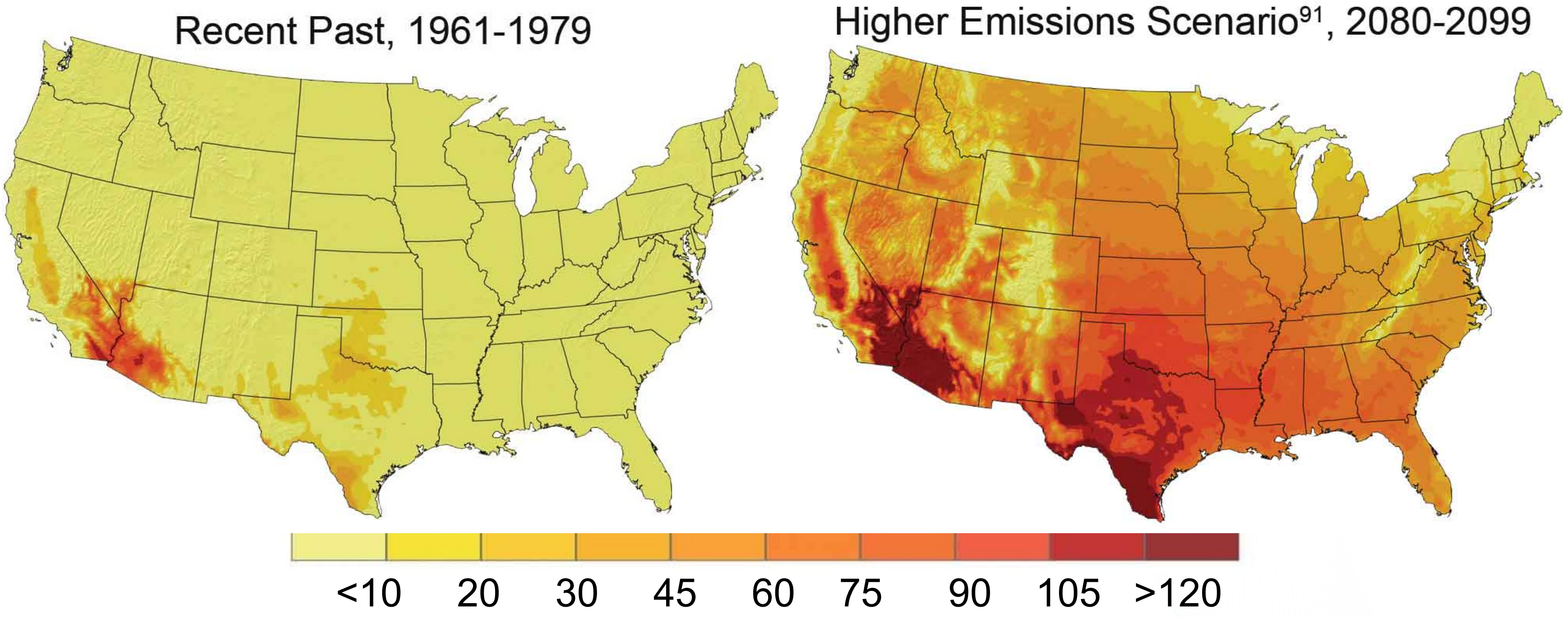
1951-1980 base

Global analysis includes
- day/night differences
- city effects

Climate variability from
- Solar irradiance
- El Nino cycle
- Arctic Oscillation
- …

http://data.giss.nasa.gov/gistemp/
(Hansen et al.)

Predicted days above 100^0 F



Much of the U.S. would go from 0 - 10 days above 100º F to 45 to 70 days per year above 100º F

Source: NOAA U.S. Global Change Research Program (climate.gov)

Carbon dioxide and climate

On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground

Svante Arrhenius

Philosophical Magazine and Journal of Science
Series 5, Volume 41, April 1896, pages 237-276.

SA was really interested in explaining ice ages…

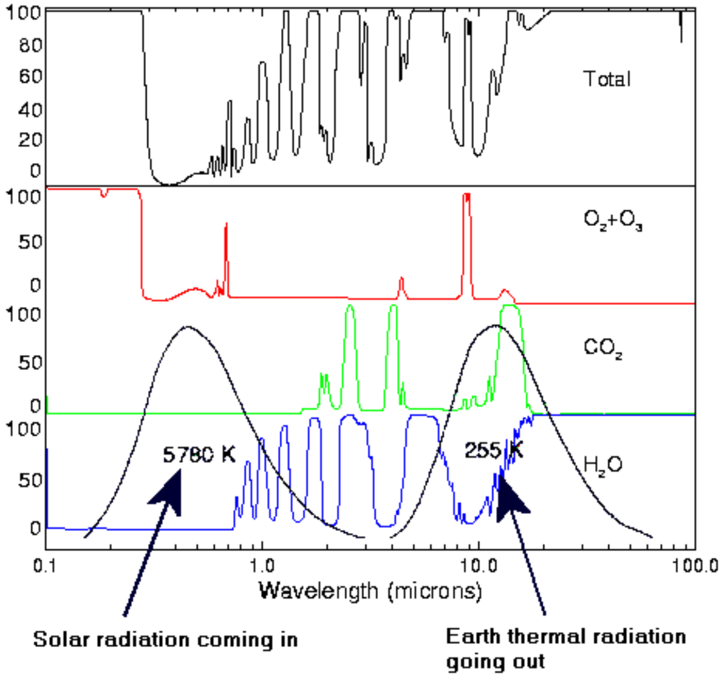
$$\Delta F = \alpha \ln\left(C \big/ C_0 \right)$$

$$\Delta T = \lambda \Delta F$$

$$\alpha = 5.35\ \mathrm{Wm}^{-2}$$

$$\lambda = 0.85 \frac{K}{\left(\mathrm{Wm}^{-2}\right)}$$

- Radiative forcing proportional to the log of the CO_2 concentration
- Proportional to T
- Doubling CO_2 means increasing temperature by about 3K (or 5.4°F)
- λ emerges from coupled-ocean-atmosphere models
- α taken from experiments



GHG sources: mainly CO_2

Sources

Energy
Electricity & power
Transportation

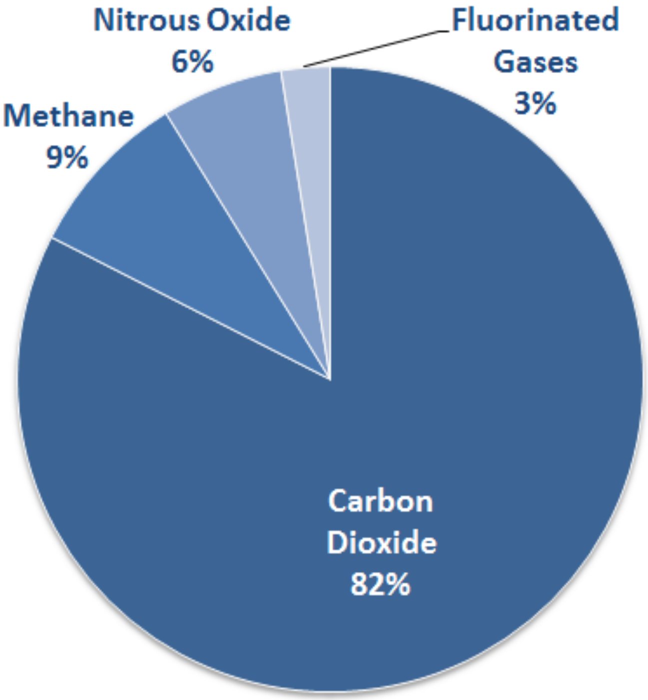
Waste

Land

Industry

Other

2012: 6.5M metric tons of CO_2 equivalent



Source:http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html

US China bilateral agreements

Copenhagen accord (January 2010)

US-China Joint Announcement on Climate Change and Clean Energy cooperation (11 November 2014)

- **Reduced US carbon footprint (reduction of 2.3-2.8% per year)**
- **China: 20% of energy needs from clean sources (800 – 1,000 GW) by 2030 (nuclear, wind, solar,…)**

29/01 2010 14:57 FAX 68505877 ⎙001

中国国家发展和改革委员会应对气候变化司
DEPARTMENT OF CLIMATE CHANGE, NATIONAL DEVELOPMENT & REFORM COMMISSION OF CHINA
No. 38, Yue Tan Nan Jie, Beijing, 100824, China, Tel: +86-10-68505862, Fax: +86-10-68505881

28 January 2010

Executive Secretary
UNFCCC Secretariat
Bonn, Germany
Fax: +49-228-8151997

Dear Mr. Yvo de Boer,

I have the honor to communicate to you the information on China's autonomous domestic mitigation actions as announced, for information to the UNFCCC Parties, as follows:

China will endeavor to lower its carbon dioxide emissions per unit of GDP by 40-45% by 2020 compared to the 2005 level, increase the share of non-fossil fuels in primary energy consumption to around 15% by 2020 and increase forest coverage by 40 million hectares and forest stock volume by 1.3 billion cubic meters by 2020 from the 2005 levels.

Please note that the above-mentioned autonomous domestic mitigation actions are voluntary in nature and will be implemented in accordance with the principles and provisions of the UNFCCC, in particular Article 4, paragraph 7.

This Communication is made in accordance with the provisions of Articles 12, paragraph 1(b), Article 12, paragraph 4 and Article 10, paragraph 2(a).

Sincerely yours,

SU Wei
Director-General
Department of Climate Change
National Development and Reform Commission of China
(National Focal Point)

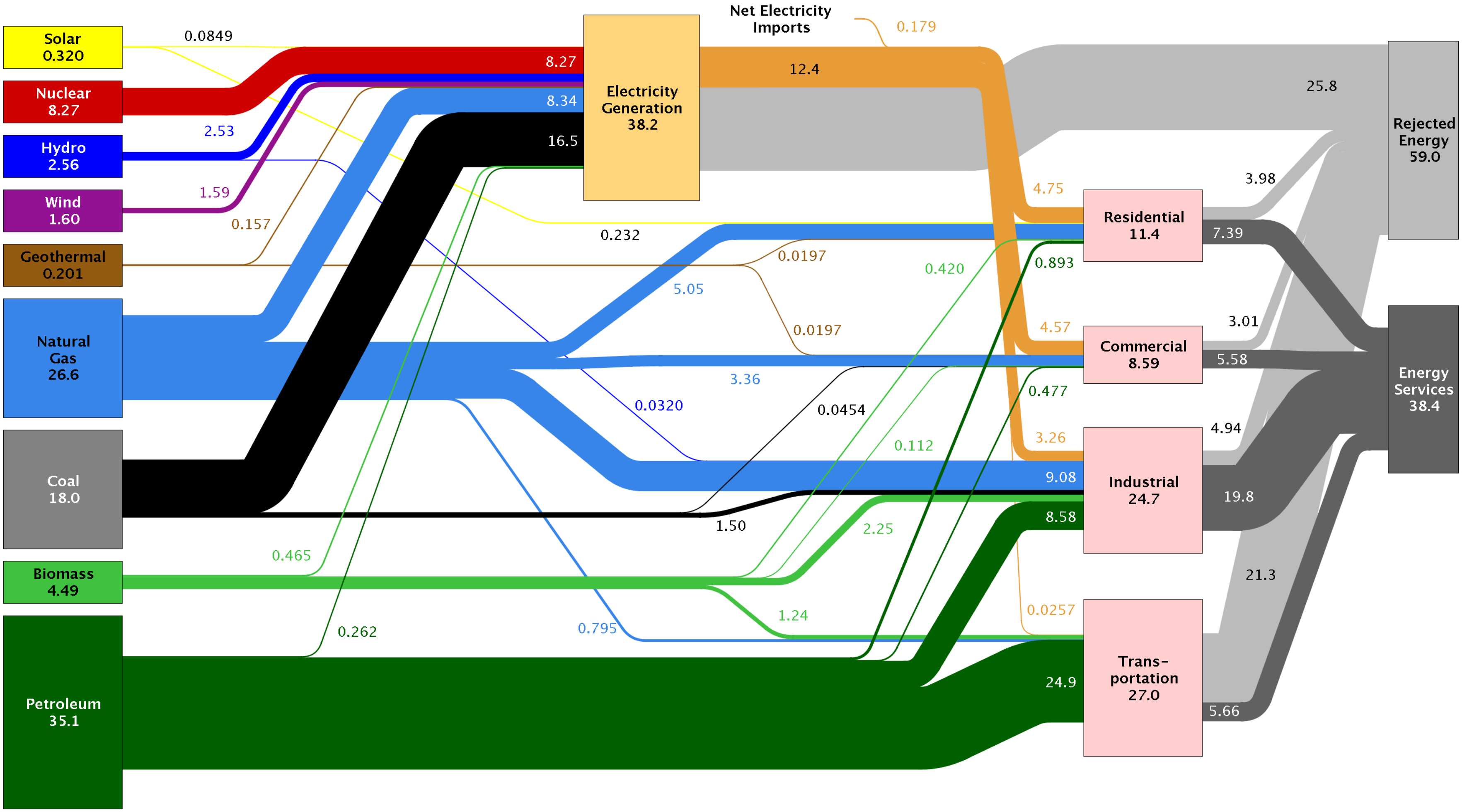
Represents a major policy driver for energy relevant materials and chemistry R&D

(http://www.chinafaqs.org/files/chinainfo/China_CPH_Accord_Submission_Letter.pdf)
http://www.whitehouse.gov/the-press-office/2014/11/11/fact-sheet-us-china-joint-announcement-climate-change-and-clean-energy-c

US Energy Production and Usage 2013 (97.4 Quads)

Estimated U.S. Energy Use in 2013: ~97.4 Quads

Lawrence Livermore National Laboratory



Source: LLNL 2014. Data is based on DOE/EIA-0035(2014-03), March, 2014. If this information or a reproduction of it is used, credit must be given to the Lawrence Livermore National Laboratory and the Department of Energy, under whose auspices the work was performed. Distributed electricity represents only retail electricity sales and does not include self-generation. EIA reports consumption of renewable resources (i.e., hydro, wind, geothermal and solar) for electricity in BTU-equivalent values by assuming a typical fossil fuel plant "heat rate." The efficiency of electricity production is calculated as the total retail electricity delivered divided by the primary energy input into electricity generation. End use efficiency is estimated as 65% for the residential and commercial sectors 80% for the industrial sector, and 21% for the transportation sector. Totals may not equal sum of components due to independent rounding. LLNL-MI-410527

1 Quad = 10^{15} BTU = 1.055×10^{18} Joule

Energy solutions through R&D

Consumption

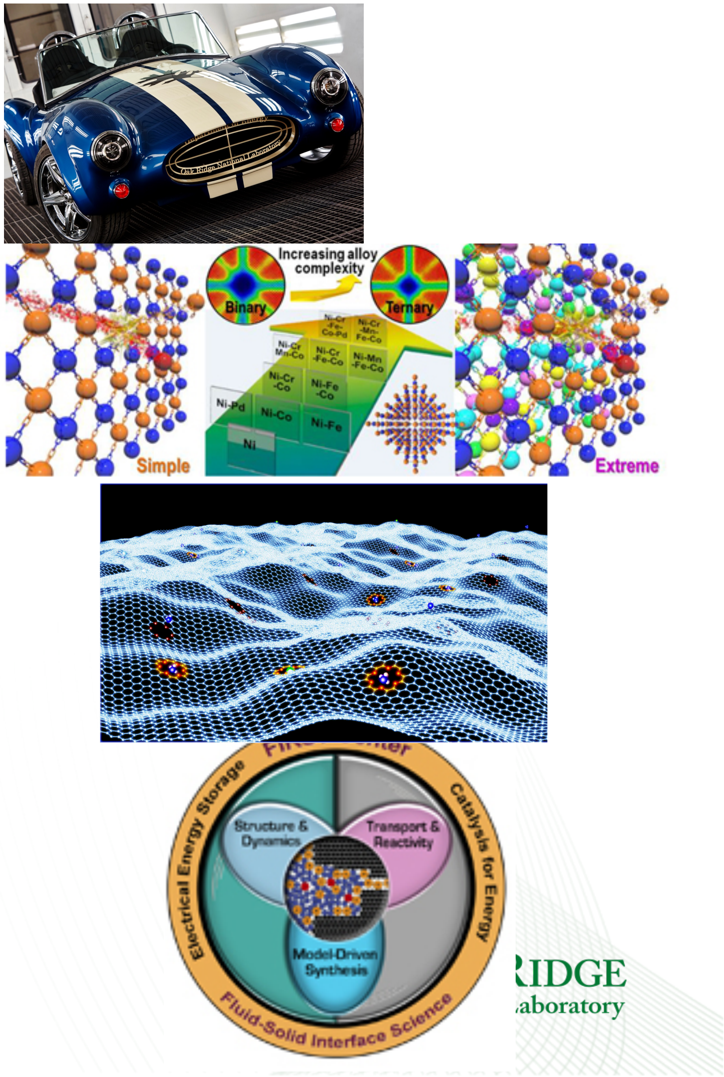
Battery technology
Advanced biofuels
Advanced manufacturing

Nuclear fuel materials
Photovoltaics
Superconductivity

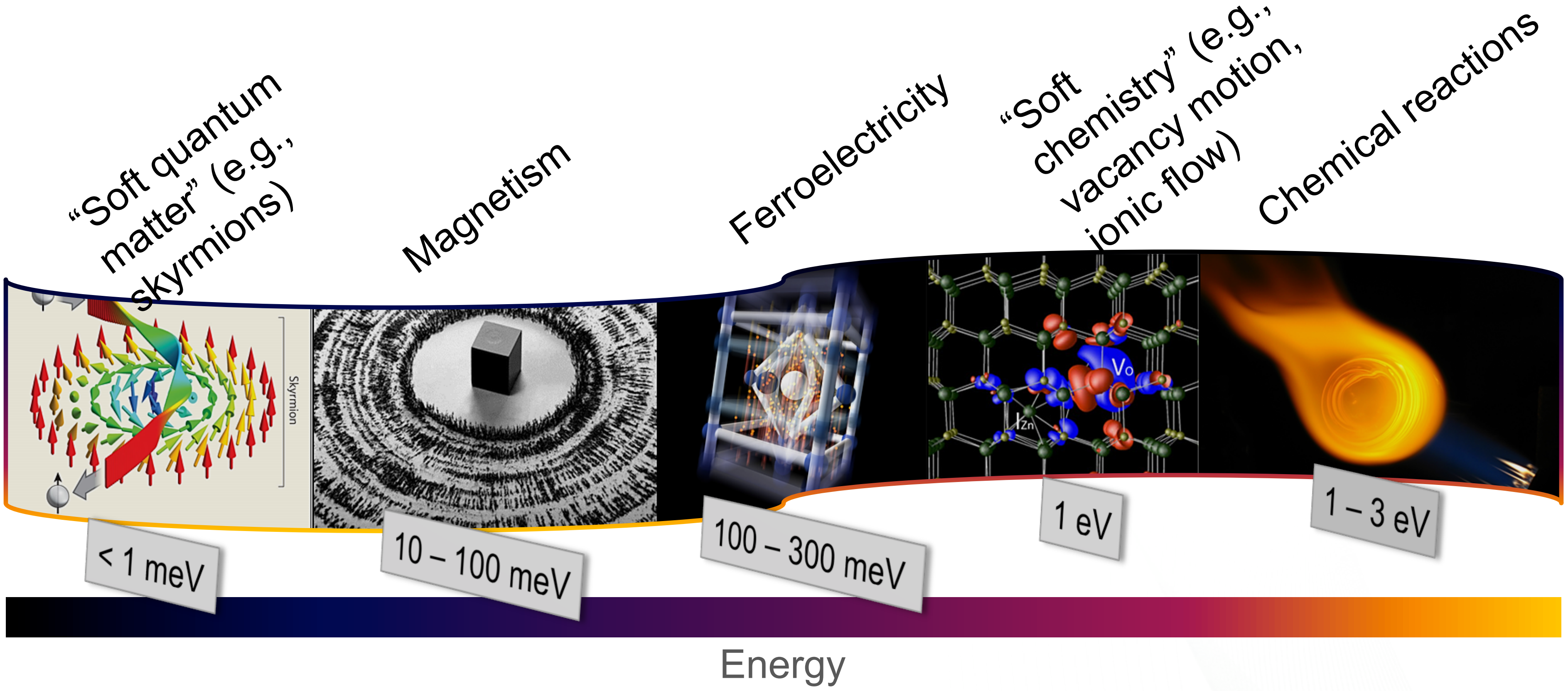
Catalysis
Separations chemistry

Theory and simulation
Materials characterization and synthesis

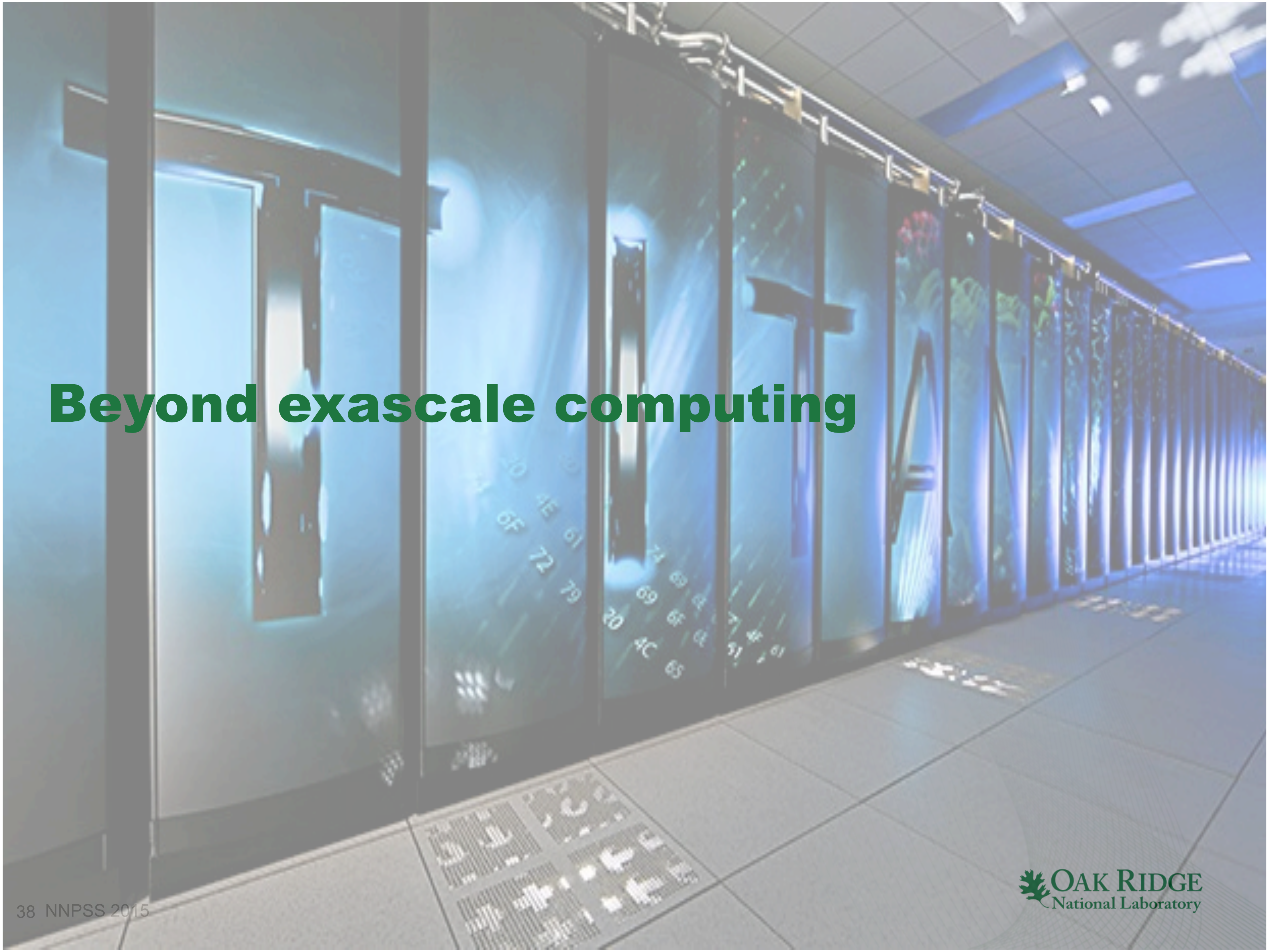
Reduced consumption through S&T



Coupled physical and chemical processes span a broad energy scale



Need to develop the appropriate tools to understand processes



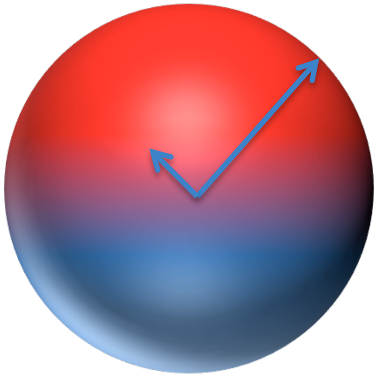
Beyond exascale computing

Science questions for quantum computing

● 1 = on

● 0 = off

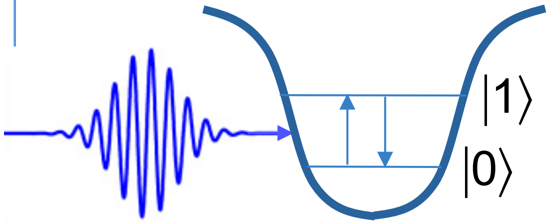
Classical: definite



$$|\Psi\rangle = a|0\rangle + b|1\rangle$$

$$|a|^2 + |b|^2 = 1$$

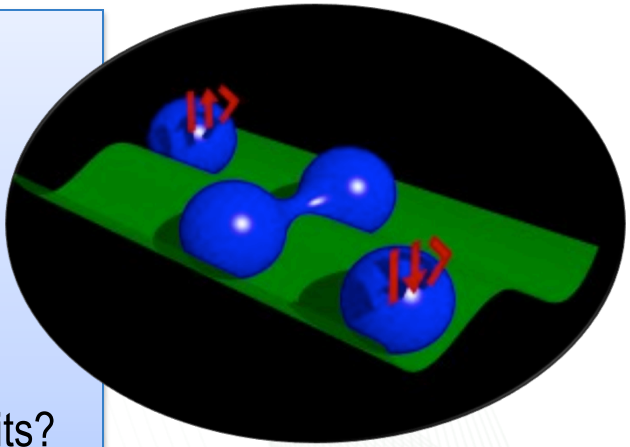
Quantum: state superposition



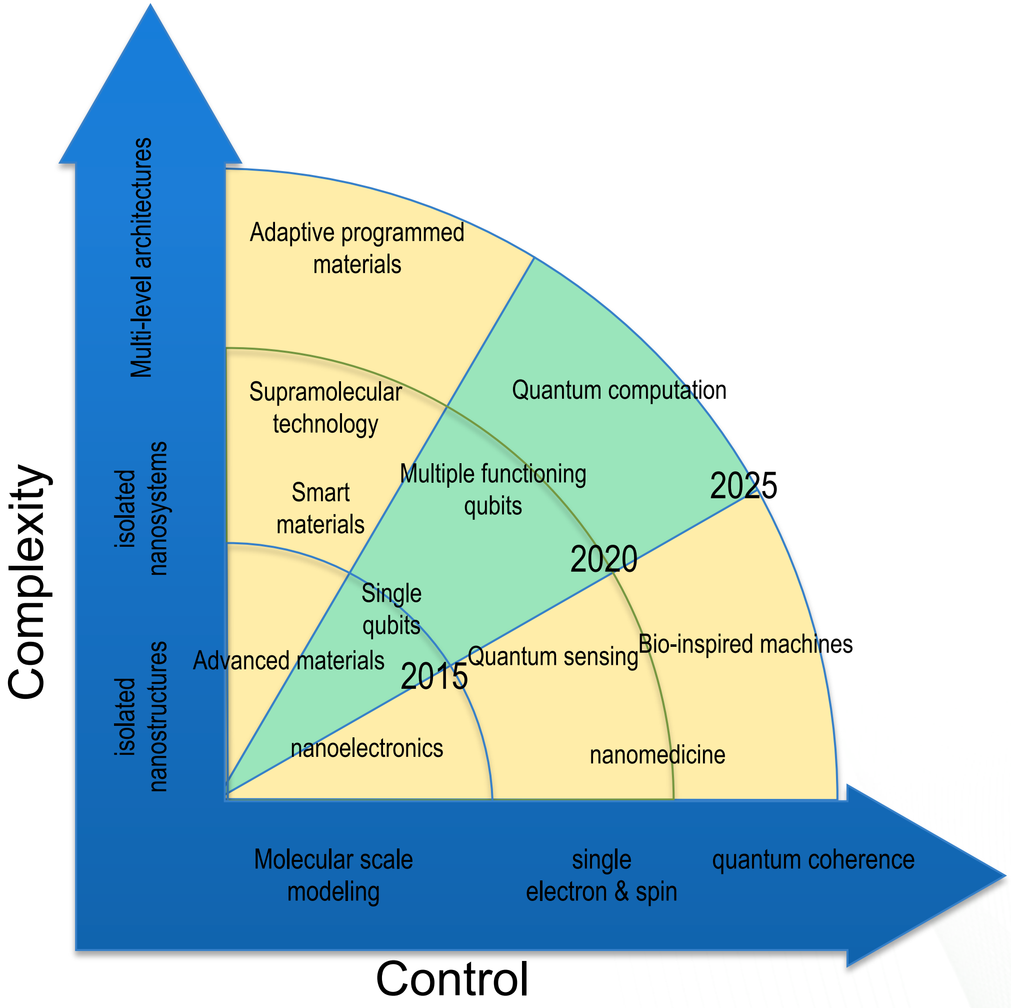
Must prepare and probe with external fields

Science questions:

- How does entanglement work across several qubits?
- How does one preserve entanglement for long times?
- How does the environment affect the entanglement?
- What quantum many-body phenomena emerge across many qubits?
- What can a set of qubits calculate (new algorithms)?



Control-complexity phase space



Classical Logic

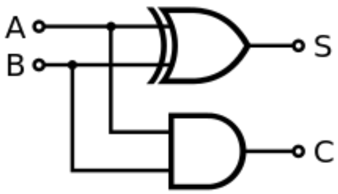
Basic gate: XOR

Input		Output
A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0

Basic gate: AND

Input		Output
A	B	A And B
0	0	0
0	1	0
1	0	0
1	1	1

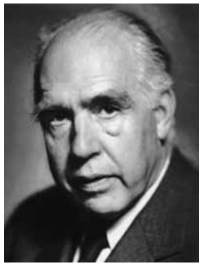
Input		Output	
A	B	C	S
0	0	0	0
1	0	0	1
0	1	0	1
1	1	1	0



Half adder

Quantum Logic

If quantum mechanics hasn't profoundly shocked you, you haven't understood it yet.
– Niels Bohr



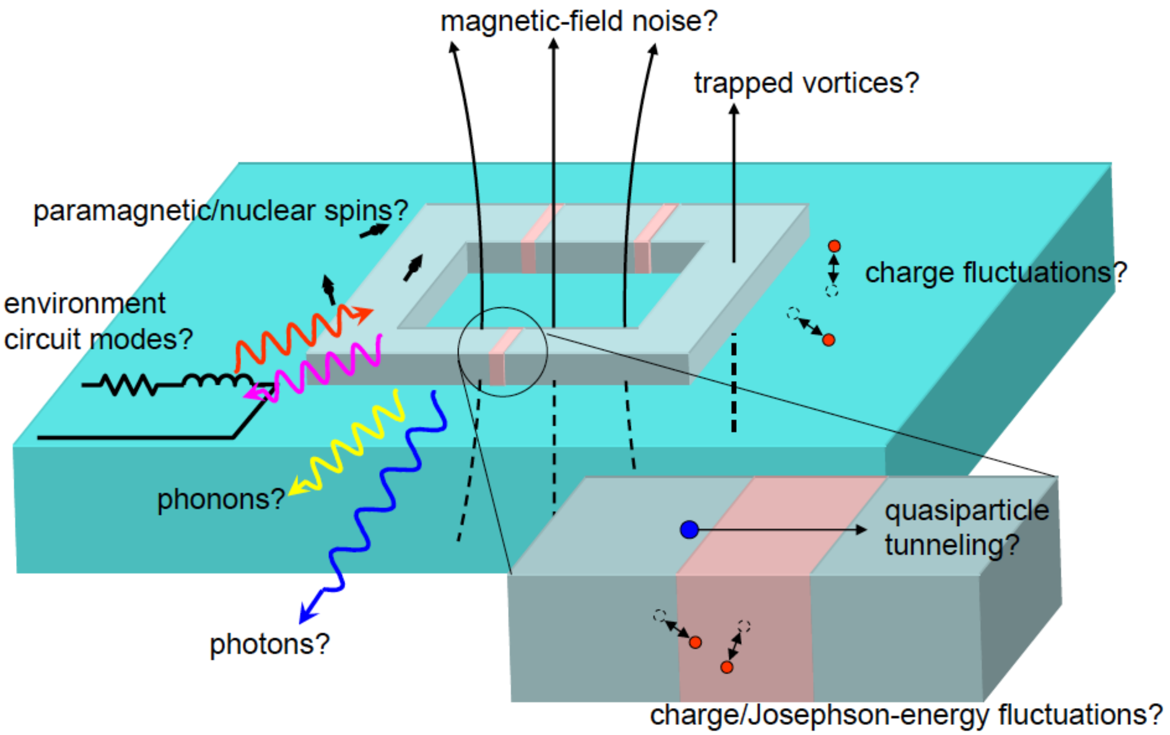
Before		After	
Control	Target	Control	Target
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

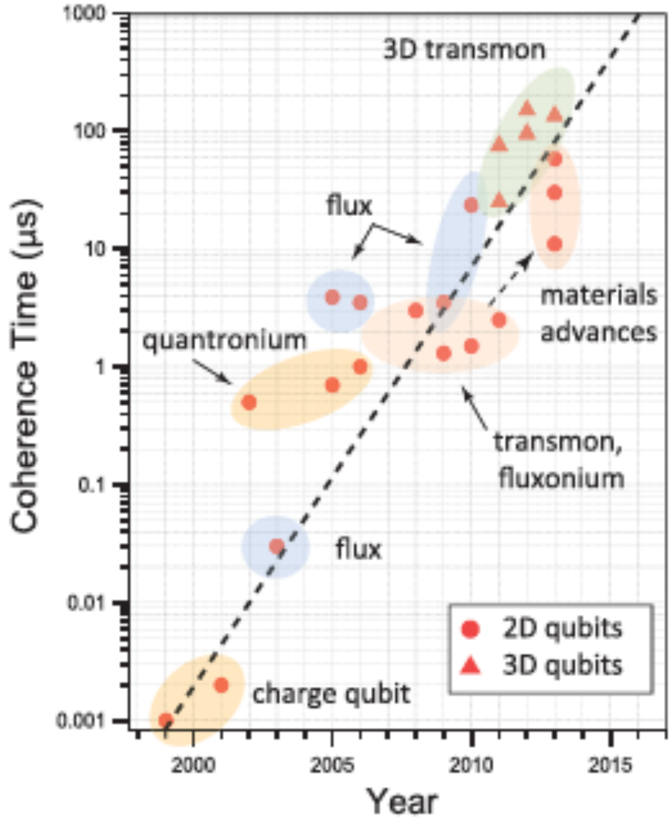
A CNOT gate flips the second bit if and only if the first bit is 1

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$$

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Qubit challenge: coherence



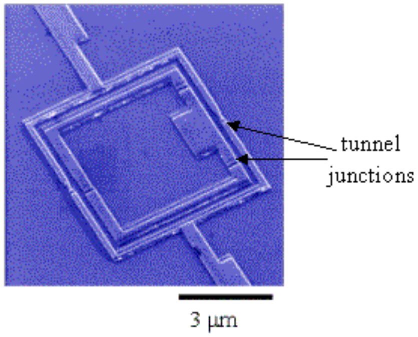


Oliver & Welander, MRS Bull. 38, 816 (2013)

Material effects cause decoherence

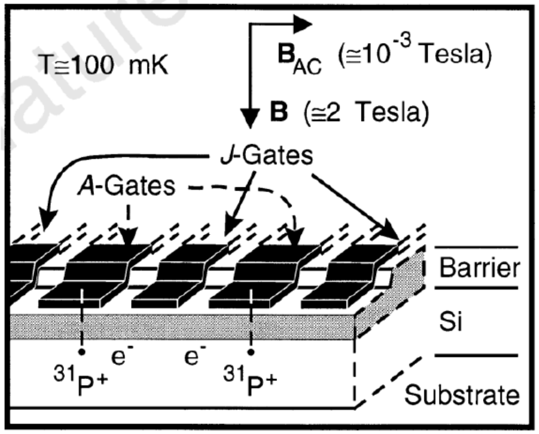
Solution:

- Reduce noise sensitivity through design, modeling and experiments
- Identify & reduce noise sources via materials and fabrication improvements

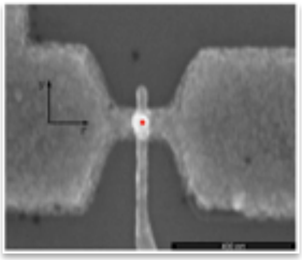


Device implementation

“Standard” Plan for Si:P-based Devices

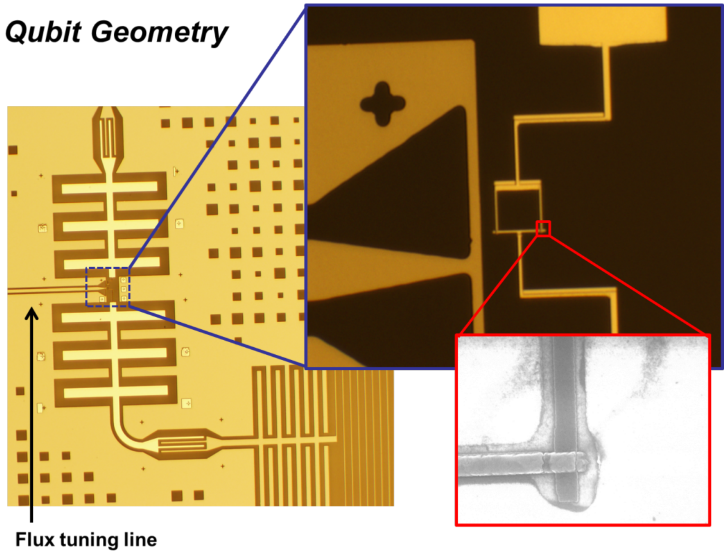


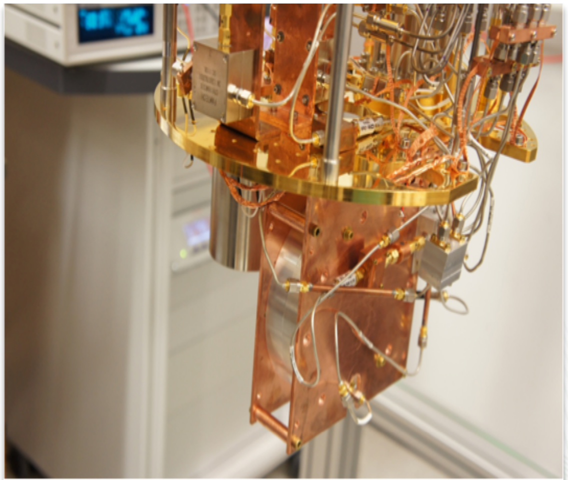
B.E. Kane, Nature **393**, 133 (1998).



Phosphorous donor from Sydney

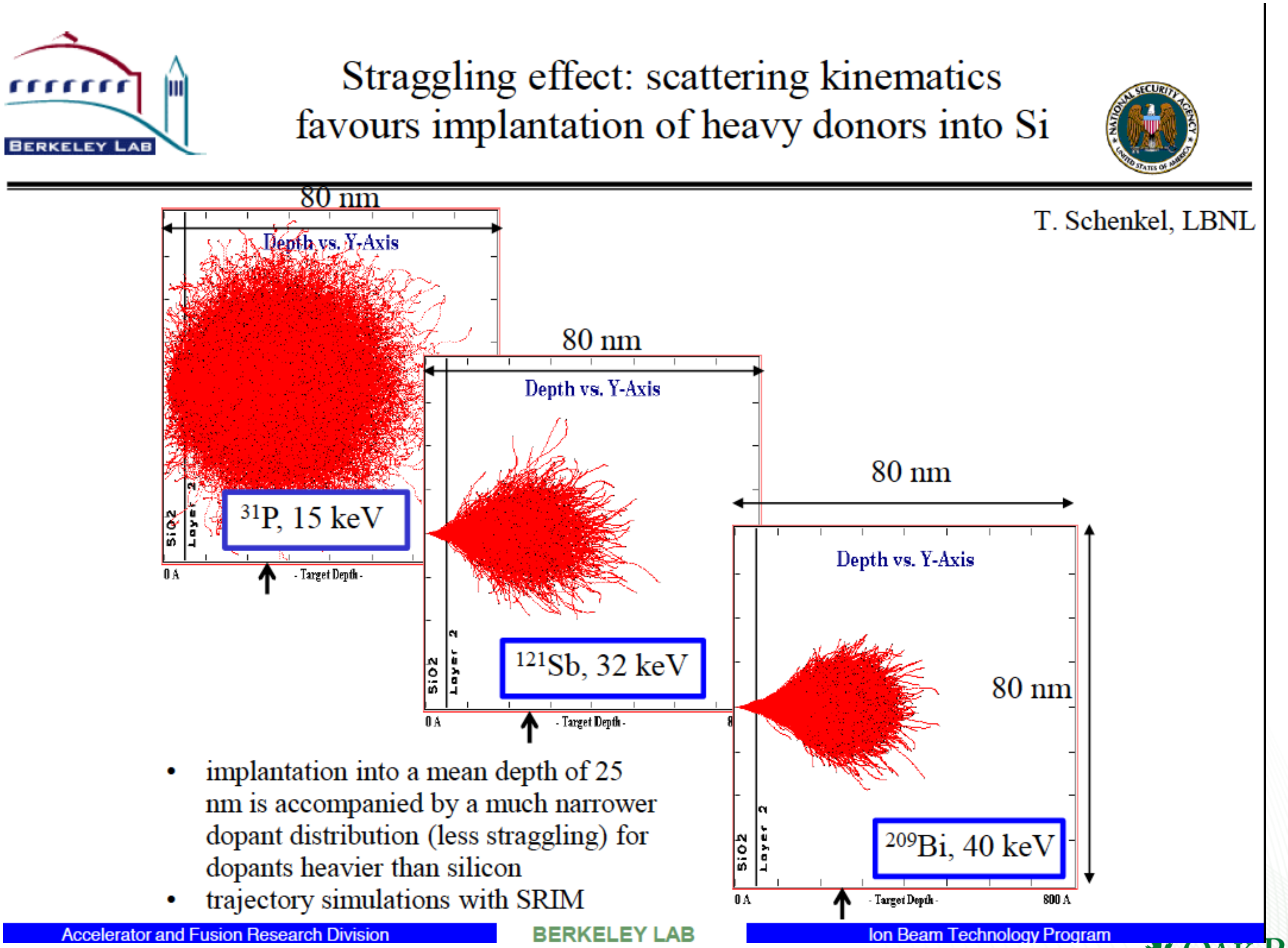
Schuster experiment (U Chicago)





Also realized in ion traps

Precision placement?



Straggling effect: scattering kinematics favours implantation of heavy donors into Si

T. Schenkel, LBNL

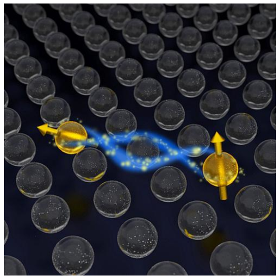
- implantation into a mean depth of 25 nm is accompanied by a much narrower dopant distribution (less straggling) for dopants heavier than silicon
- trajectory simulations with SRIM

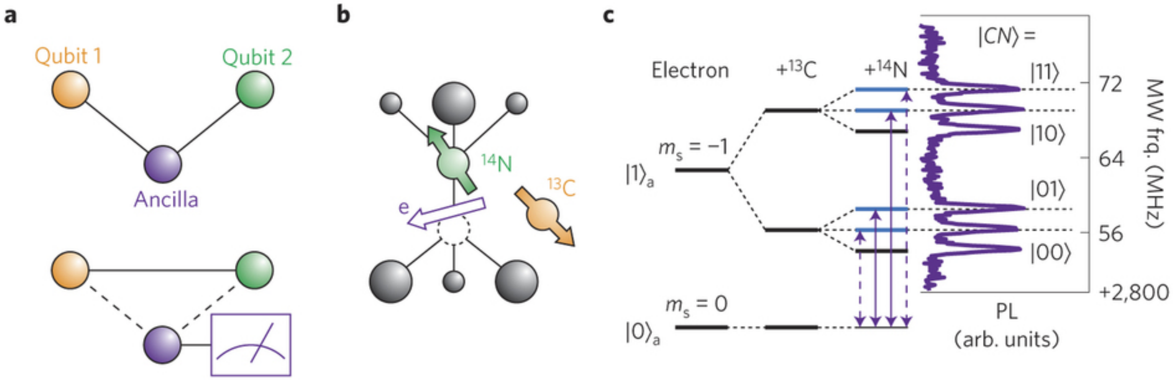
Accelerator and Fusion Research Division

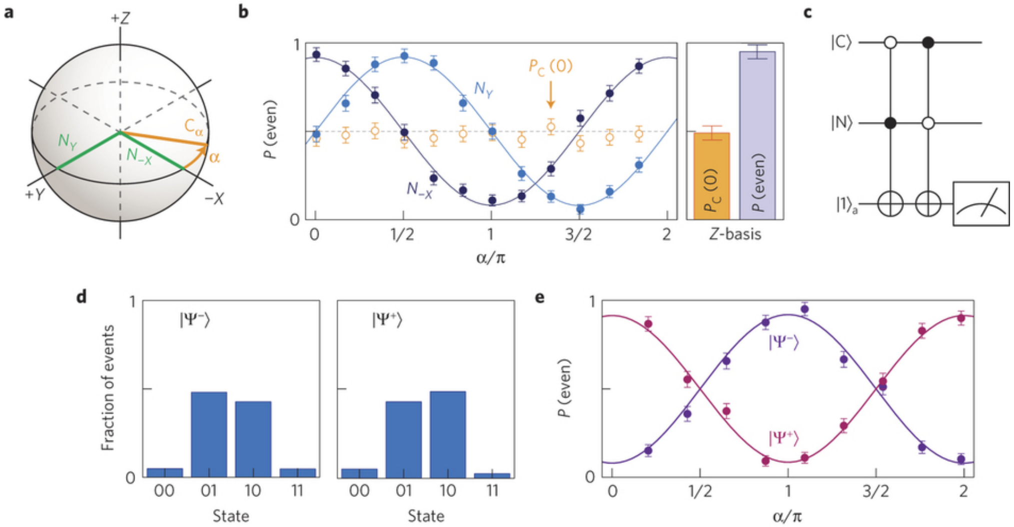
BERKELEY LAB

Ion Beam Technology Program

Quantum measurement in diamond



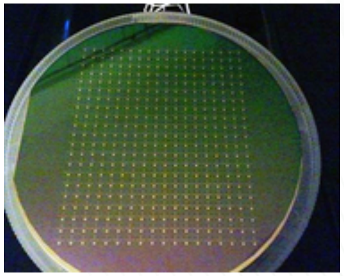




- State preparation
- State manipulation (rf-field)
- Setting up a Toffoli gate
- Measurement of the entangled state

Everything about 'quantum computing' is time dependent

Demonstration of entanglement by measurement of solid state qubits, Pfaff, Nature Physics 9, 29 (2013)



ANL public-private partnership with AKHAN Semiconductor

OAK RIDGE National Laboratory

Science and engineering questions

- **Materials**

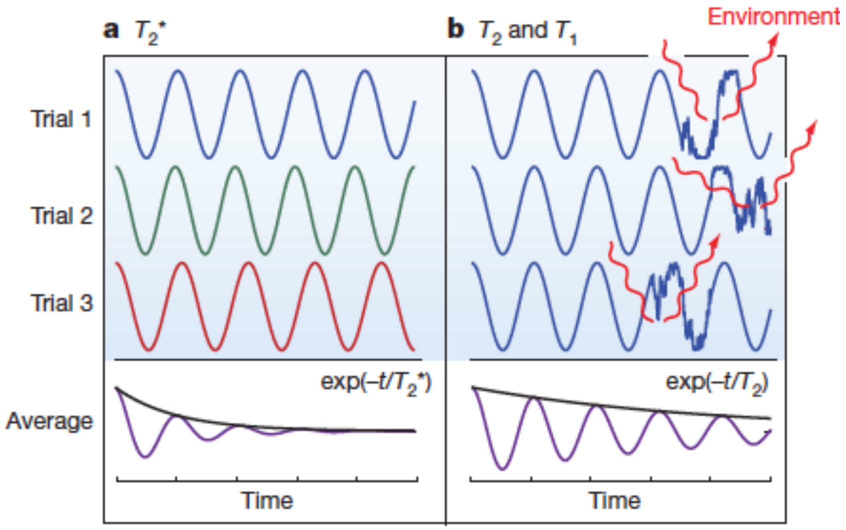
  What are the materials properties that must be controlled in order to increase coherence (in time and space)?

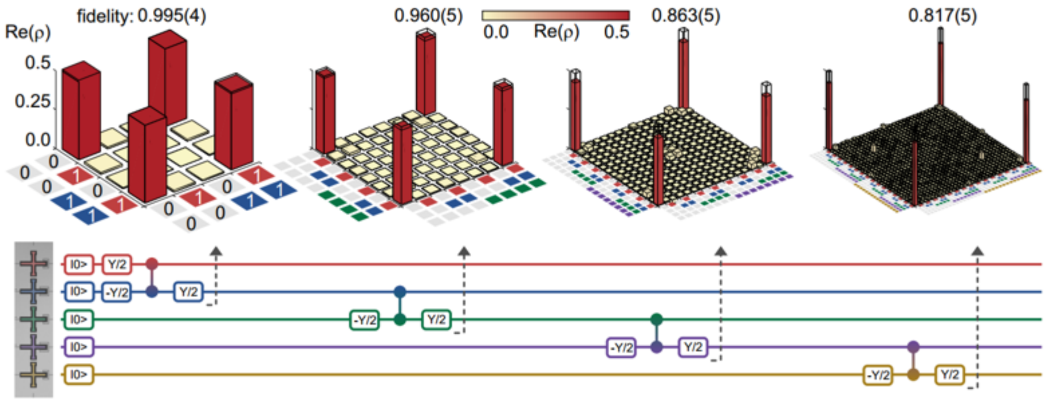
- **Qubits**

  How do we *reliably* make a few entangled qubits?

- **Computing models**

  What is the computing model for coupling quantum and classical computing?





5 qubit reliability in ion traps = 82%
Nature 508, 500 (2014)

Discussion and conclusion

- We have great tools for scientific discovery
  - Experimental facilities: RHIC, TJLab, FRIB, ANL, LHC, SNS,…
  - Computational (theoretical) facilities: Leadership computing, NERSC
- We have great problems to solve
  - QCD, quantum many-body problems, astrophysics, beyond SM
  - Significant discovery potential
- You can and do make a difference