High-Impact Science on NERSC's Cori:

A KNL success story

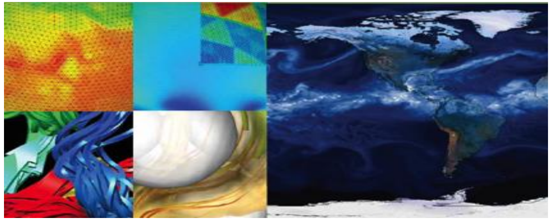
Richard Gerber

NERSC Senior Science Advisor
High Performance Computing Department Head

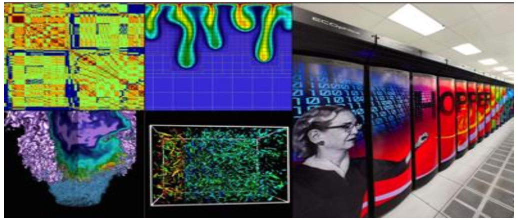
NERSC: the Mission HPC Facility for DOE Office of Science Research

U.S. DEPARTMENT OF ENERGY | Office of Science

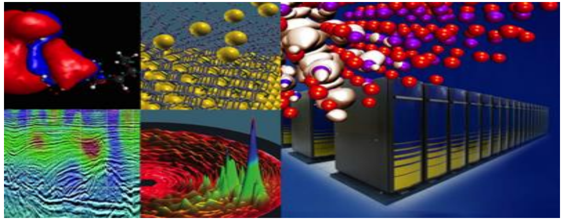
Largest funder of physical science research in the U.S.



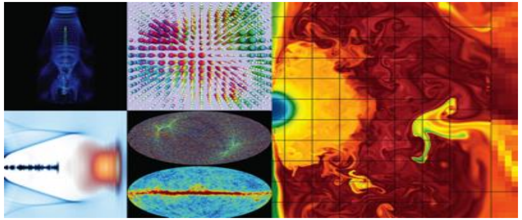
Bio Energy, Environment



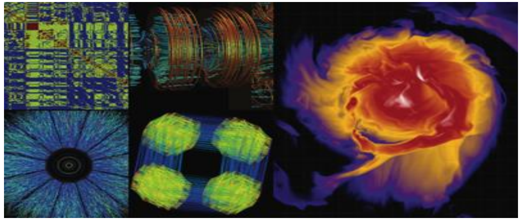
Computing



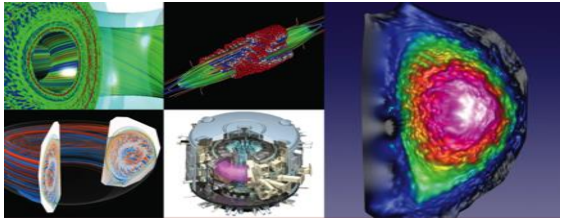
Materials, Chemistry, Geophysics



Particle Physics, Astrophysics



Nuclear Physics



Fusion Energy, Plasma Physics

6,000 users, 700 projects, 700 codes, 48 states, 40 countries, universities & national labs

Cray XC40 system with 9,600+ Intel Knights Landing compute nodes

68 core KNL / 272 HW threads

96 GB DRAM / 16 GB MCDRAM

Support the entire Office of Science research community

Begin to transition workload to energy efficient architectures for NERSC users

Data Intensive Science Support

2K Haswell processor nodes (Phase 1)

NVRAM Burst Buffer 1.5 PB, 1.5 TB/sec

30 PB of disk, >700 GB/sec I/O bandwidth

Integrated with Cori Haswell nodes on Aries network for data / simulation / analysis on one system

Software defined networking

Innovative job scheduling

Key Differences

Edison (Xeon Ivy Bridge) 2014 Production System	Cori (Xeon Phi KNL) 2017 Production System
5,500 nodes	9,600 nodes
12 cores per socket	68 cores per socket
24 hardware threads	272 hardware threads
2.4 GHz	1.4 GHz
8 DP ops per cycle	32 DP ops cycle
64 GB DDR @ 100 GB/s node	96 GB DDR @ 120 GB/s 16 GB MCDRAM @ 450 GB/s

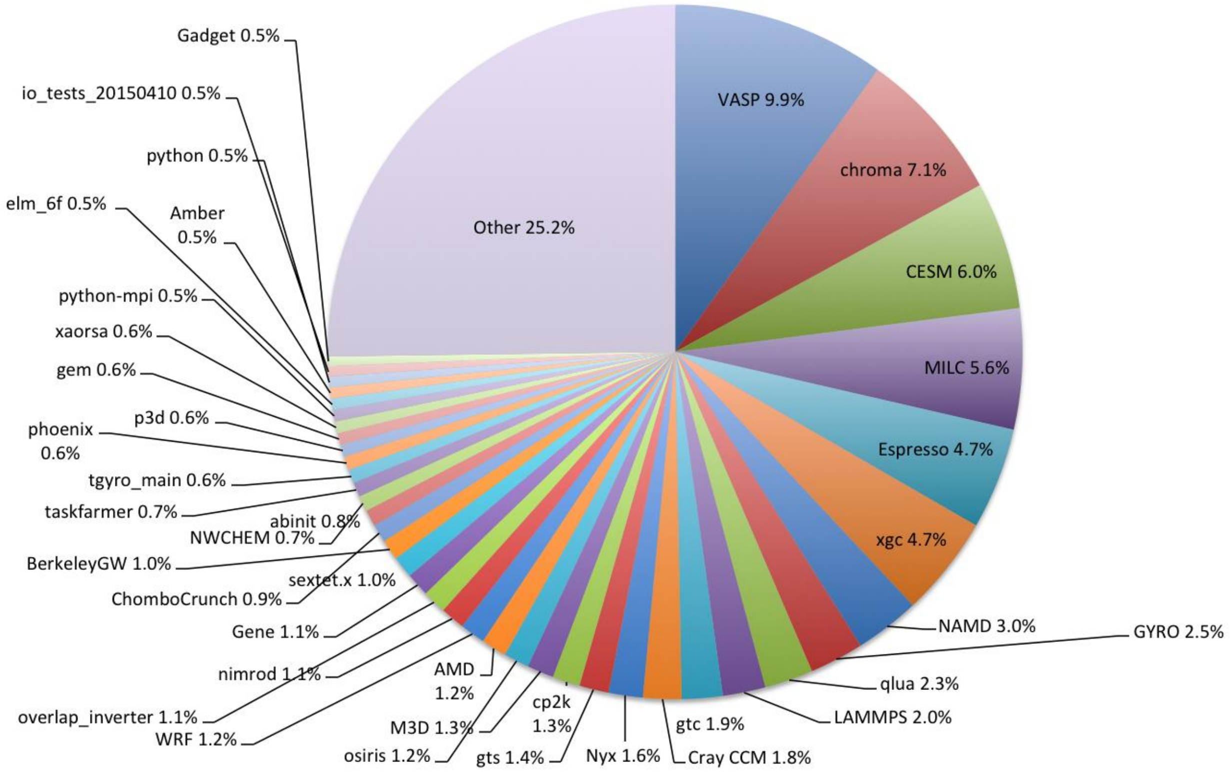
Optimization targets: OpenMP Threading, Vectors, Data management for MCDRAM

Because of these differences, we knew codes would have to adapt.

NERSC’s KNL Challenge:

Enable NERSC’s diverse community of 7,000 users, 750 projects, and 700 codes to run on Cori’s Intel Xeon Phi Knights Landing processors at high performance

NERSC Workload 2015



NERSC runs 700+ codes from 700+ projects

Top 20 codes make up more than 50% of hours used

If NERSC could get those codes – or proxies – running on KNL, Cori would be productive for science

NERSC Exascale Scientific Application Program (NESAP)

Goal: Prepare DOE Office of Science users for Cori's manycore CPUs

Partner with ~20 application teams from all 6 DOE science offices and apply lessons learned to broad NERSC user community

IXPUG
THE INTEL XEON PHI USERS GROUP

Community Confluence

- Close interactions with vendors
- Developer Workshops
- Early engagement with code teams
- Postdoc Program
- Engage in the Broad Community
- Training and online modules
- Early access to KNL

NESAP Optimization Strategy and Goals

We're primarily working with existing codes to get them ready for Cori

Goals

- Standard constructs for portability and maintainability
- Incorporate optimizations into code base by working directly with developers
- Collaborate closely with community to leverage expertise and expand NERSC influence and relevance

Strategy: Focus first on single-node optimization

- Enable fine-grained parallelism on light-weight cores via OpenMP
- Exploit dual 512b vector units
- Exploit 5X memory bandwidth due to MCDRAM by managing data access

KNL One Year Ago

ISC 2016

Lots of Excitement

Great reports of optimization success

KNL white boxes with preproduction chips

Zero science runs with full applications at scale

KNL at the One Year Mark (at NERSC)

NERSC

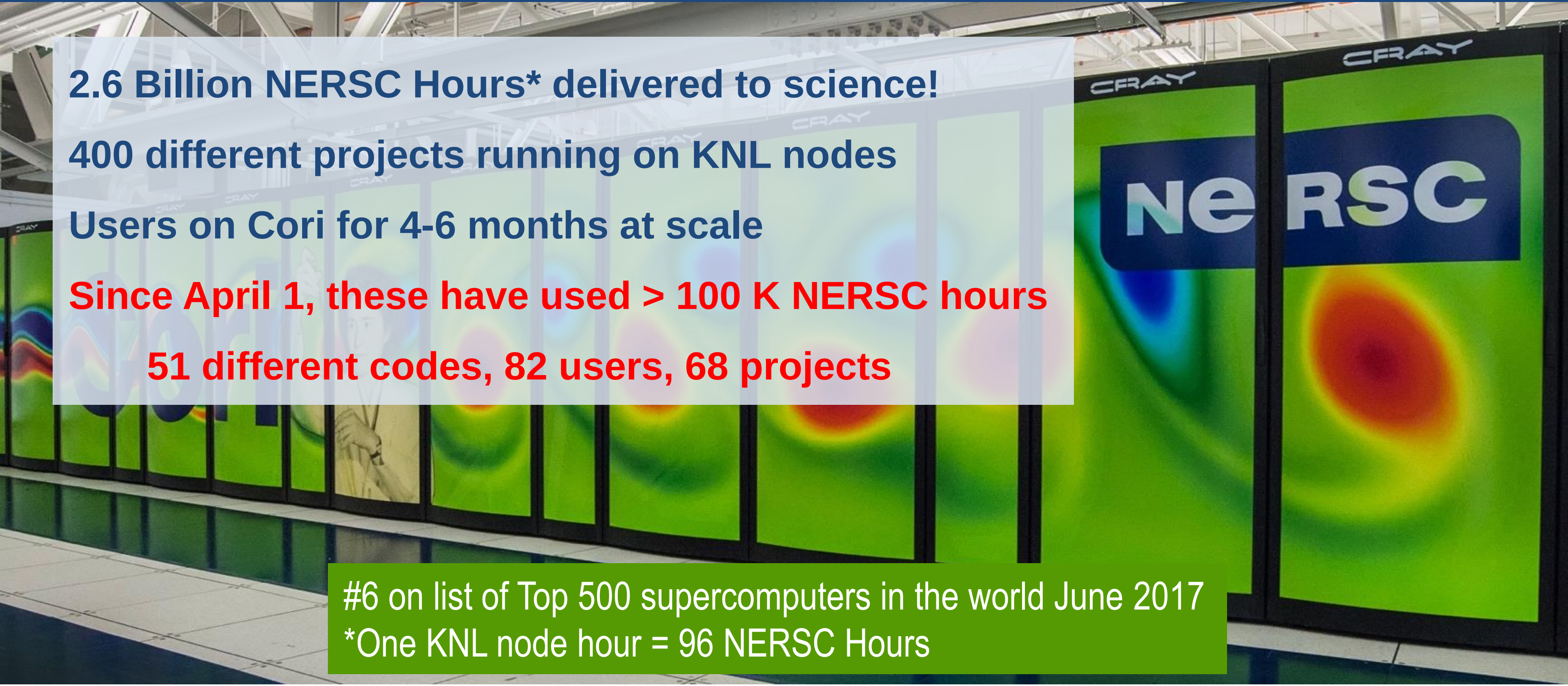
2.6 Billion NERSC Hours* delivered to science!

400 different projects running on KNL nodes

Users on Cori for 4-6 months at scale

Since April 1, these have used > 100 K NERSC hours

- **51 different codes, 82 users, 68 projects**



#6 on list of Top 500 supercomputers in the world June 2017
*One KNL node hour = 96 NERSC Hours

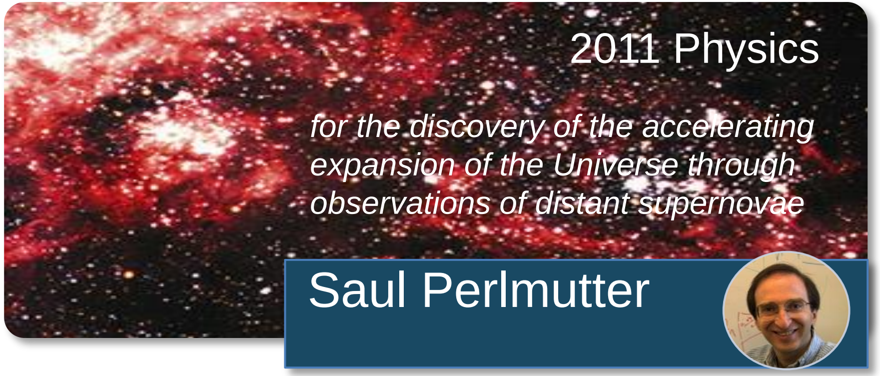
Scientific Productivity

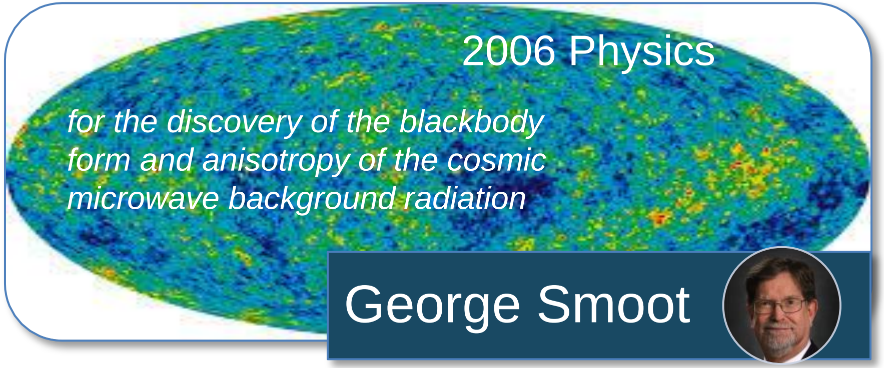
With 2.4 billion NERSC Hours in 2016, NERSC users produced more than 2,000 referred scientific publications.

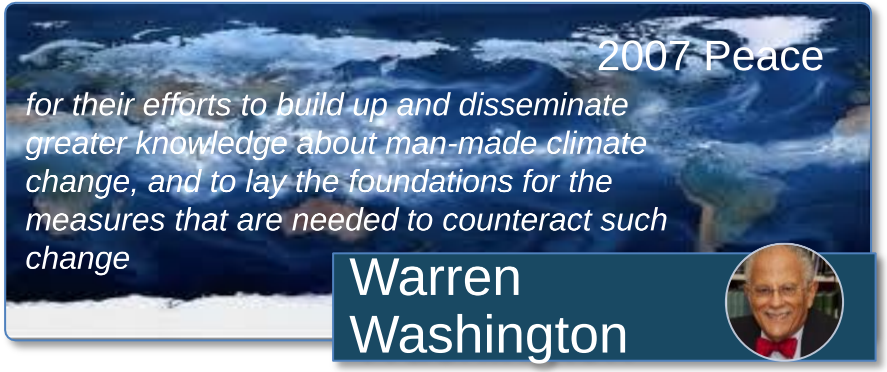
As long as the broad community can use KNL, imagine what will they do with Cori's 4.8 billion NERSC Hours per year?

2,277 refereed publications in 2016

More Nobel Prizes?







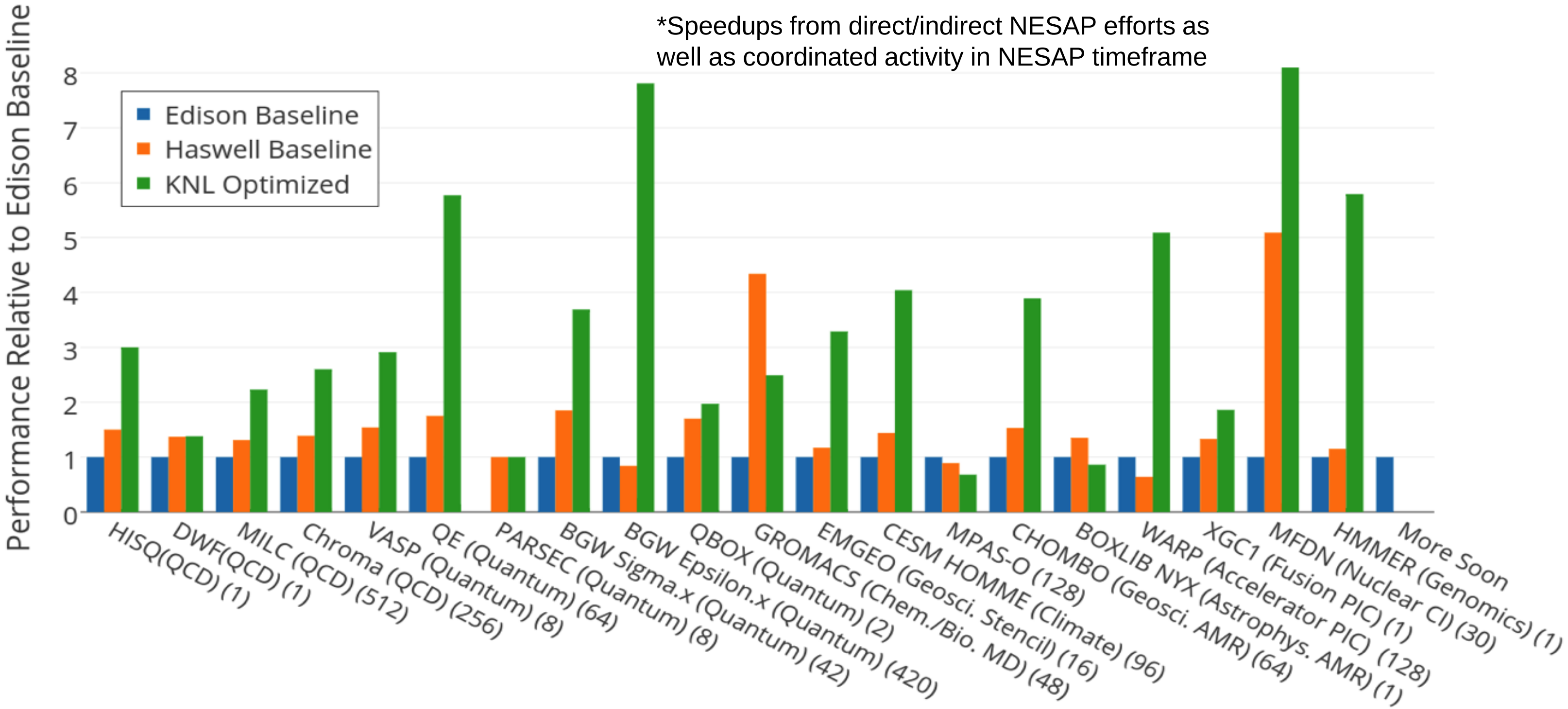
IXPUG
THE INTEL XEON PHI USERS GROUP
Community Confluence

This wouldn't be possible without you!

Results that follow are from NERSC, but typical of IXPUG results

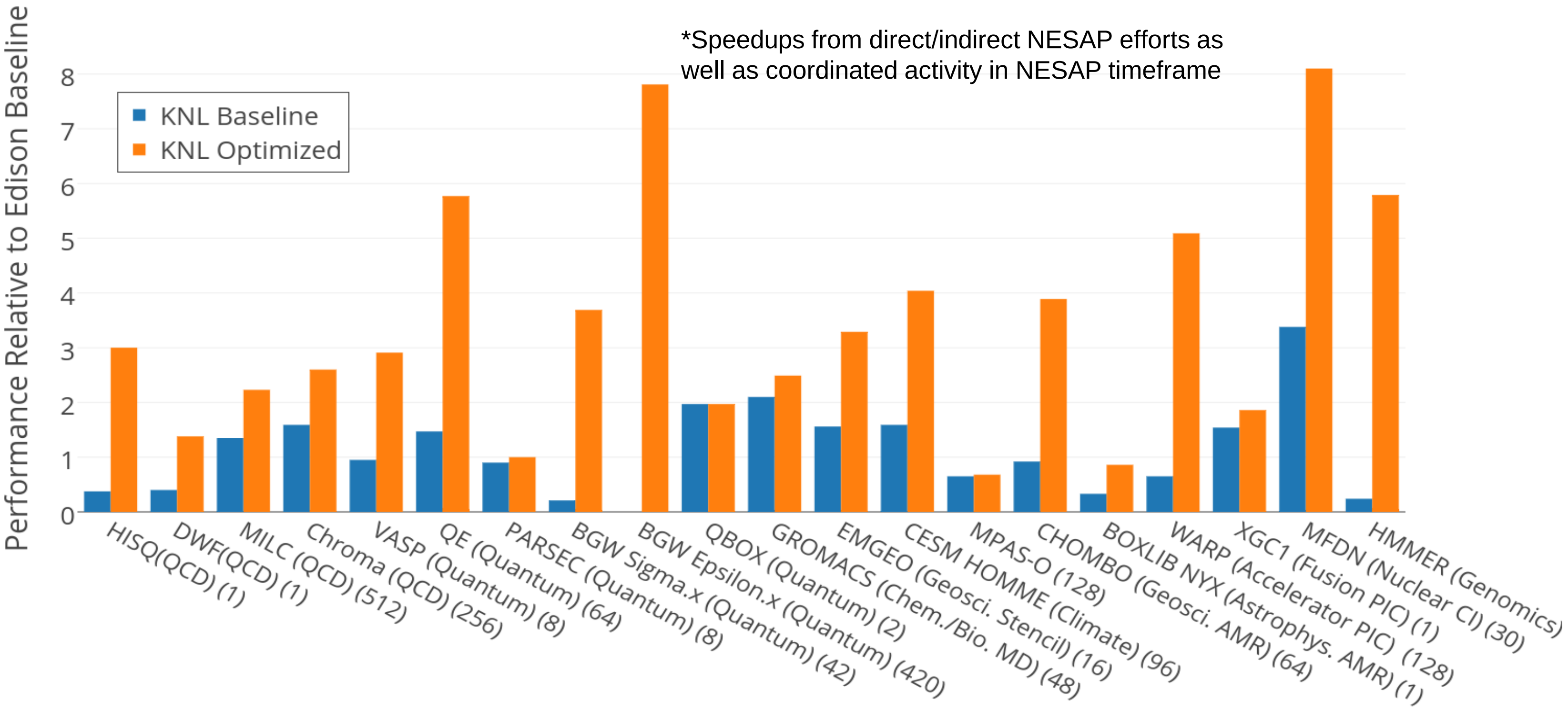
KNL Improvements vs. “Business as Usual”

*Speedups from direct/indirect NESAP efforts as well as coordinated activity in NESAP timeframe



NESAP Code Improvements on KNL

*Speedups from direct/indirect NESAP efforts as well as coordinated activity in NESAP timeframe



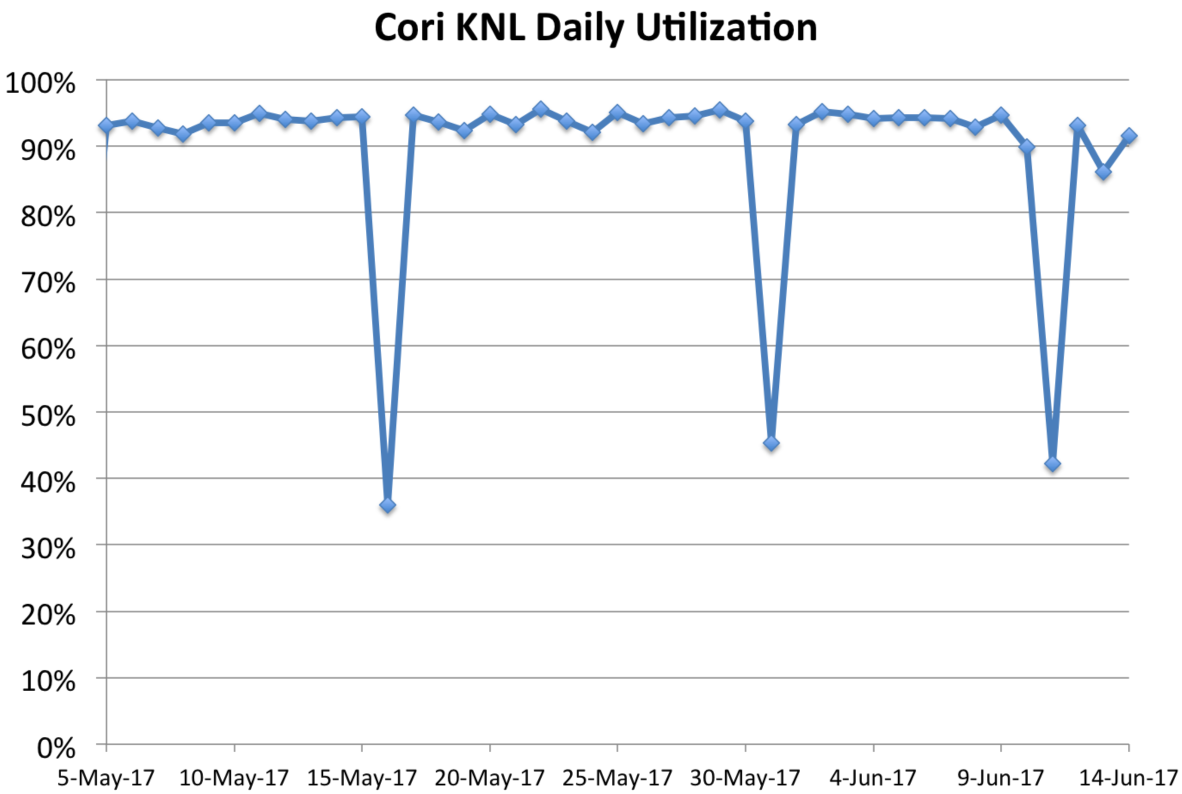
Average NESAP Code Performance (per node)

$$\frac{\text{KNL Optimized}}{\text{Haswell Optimized}} \approx 1.2$$

& codes are on the path to exascale

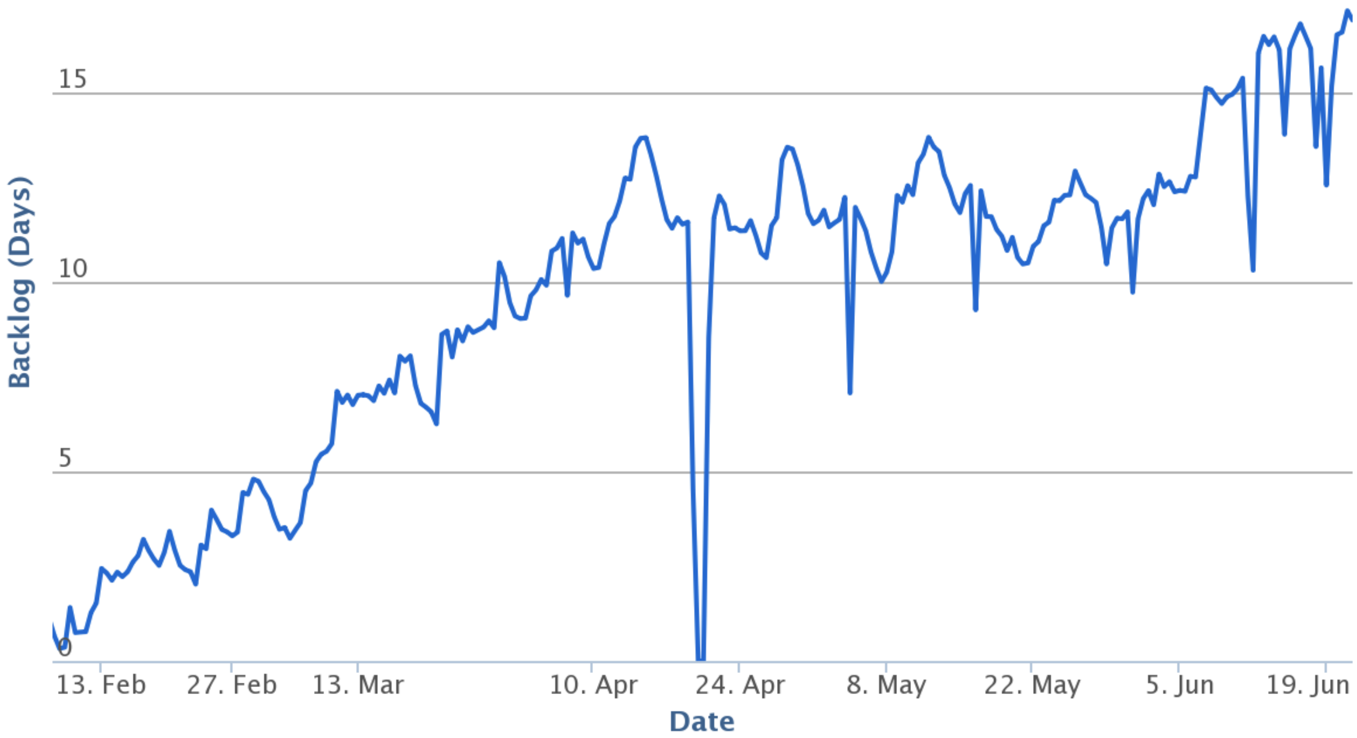
Users have jumped on KNL

Contrary to our concerns, utilization is effectively 100%



Cori KNL Backlog

Queues are completely full.

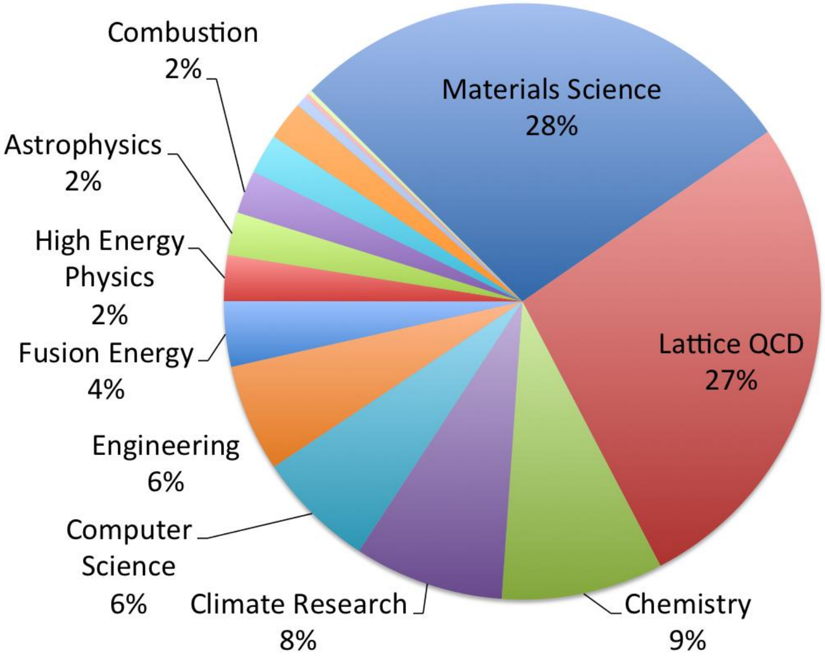


KNL Science 2017

All areas of science are using KNL

We were worried that MatSci and Chemistry users would be slow adopters, but that's not the case.

Cori KNL Hours Used Jan-Jun 2017



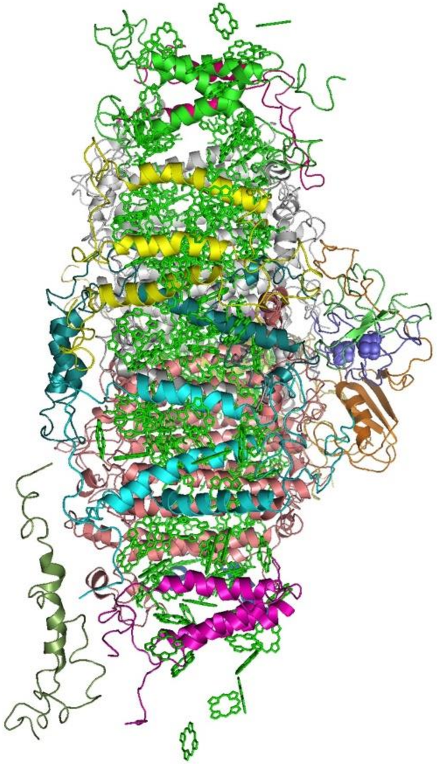
KNL Impact on Science

NERSC’s Top KNL Users in Each Science Area

Molecular Dynamics Simulations of Protein Dynamics and Lignocellulosic Biomass

The Photosystem I (PS1) is a membrane protein complex that captures solar energy and stores it in the complex.

It can couple with H2ase to produce Hydrogen from sunlight and water. Using MD simulations to study PS1/H2ase interactions in detail and learn how to produce a clean fuel.



PI: Jeremy Smith, Oak Ridge National Laboratory

NERSC Hours on KNL: 54 million

Max Concurrency: 1,080 nodes (73K cores)

Code: GROMACS

NERSC contact: Zhengji Zhao

NESAP

Accelerated Climate Modeling for Energy

Water Cycle: How will the water cycle evolve in a warmer climate?

Biogeochemistry: How will terrestrial and coastal ecosystems drive natural sources & sinks of CO_2 and CH_4 in a warmer world?

Cryosphere system: How will more extreme storms enhance the coastal impacts of sea level rise?

PI: Lai-Yung Ruby Leung, Pacific NW National Lab

NERSC Hours on KNL: 109 million

Max Concurrency: 8,192 nodes (557K cores)

Code: ACME

NERSC contact: Helen He

NESAP

First-Principles Catalyst Design for Environmentally Benign Energy Production

Designing new, cheaper, better, and environmentally benign catalysts for production and chemical utilization of hydrogen, for production of hydrocarbon fuels, and for low temperature fuel cells.

PI: Manos Mavrikakis, U. of Wisconsin, Madison

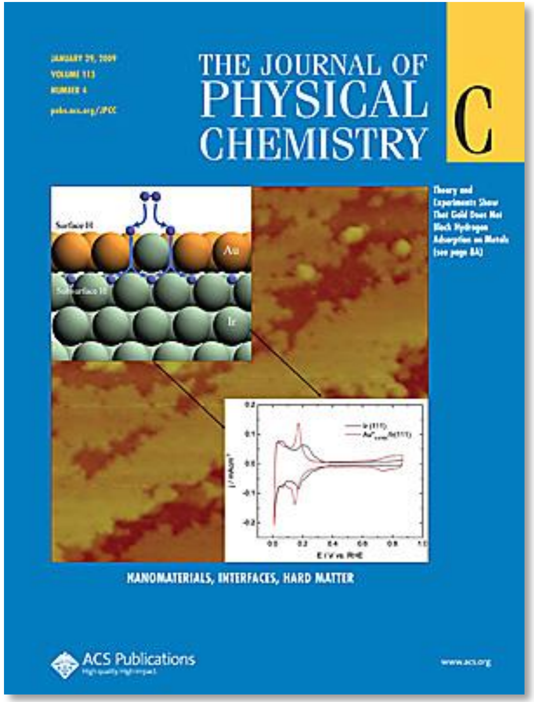
NERSC Hours on KNL: 201 million

Max Concurrency: 64 nodes (4.3K cores)

Code: VASP

NERSC contact: Zhengji Zhao

NESAP



Large Scale 3D Geophyscial Inversion & Imaging

Imaging subsurface geophysical properties in 3D and relating these properties to critical geological processes relevant to energy exploration and carbon sequestration.

PI: Jeff Newman, Berkeley Lab

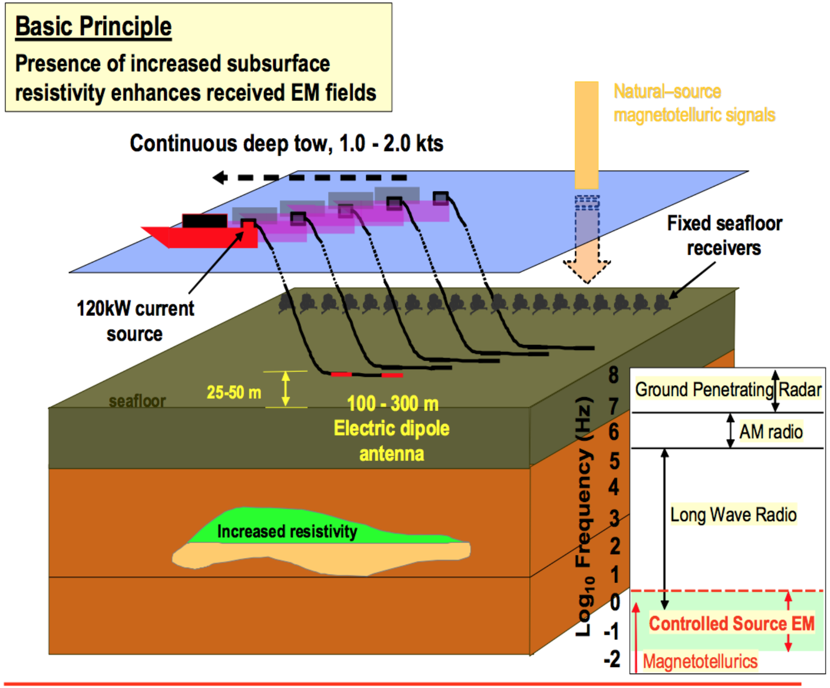
NERSC Hours on KNL: 38 million

Max Concurrency: 520 nodes (35K cores)

Code: EMGeo

NERSC contact: Thorsten Kurth

NESAP



The Materials Genome

"In the lab" materials synthesis and discovery can take 20 years to get to market.

The materials project is accelerating the way materials discovery is done by creating a high-throughput computing environment together with a searchable, interactive database of computed materials properties.

MATERIALS PROJECT

PI: Kristin Ceder-Persson, Berkeley Lab

NERSC Hours on KNL: 109 million

Max Concurrency: 8 nodes (544 cores, 150K runs)

Code: VASP

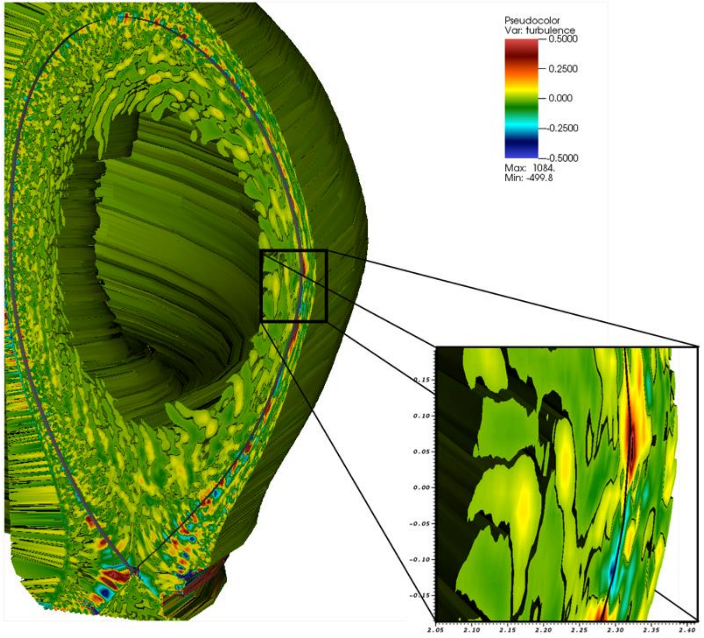
NERSC contact: Zhengji Zhao

NESAP

Center for Edge Physics Simulation: SciDAC-3 Center

Understanding multi-scale, non-thermal edge physics in fusion reactors.

Edge physics crucially affects both the core energy production and the erosion rate of edge materials. Both are keys to creating clean energy from fusion reactions.



PI: C.S. Chang, Princeton Plasma Physics Lab

NERSC Hours on KNL: 71 million

Max Concurrency: 4,096 nodes (279K cores)

Code: XGC1

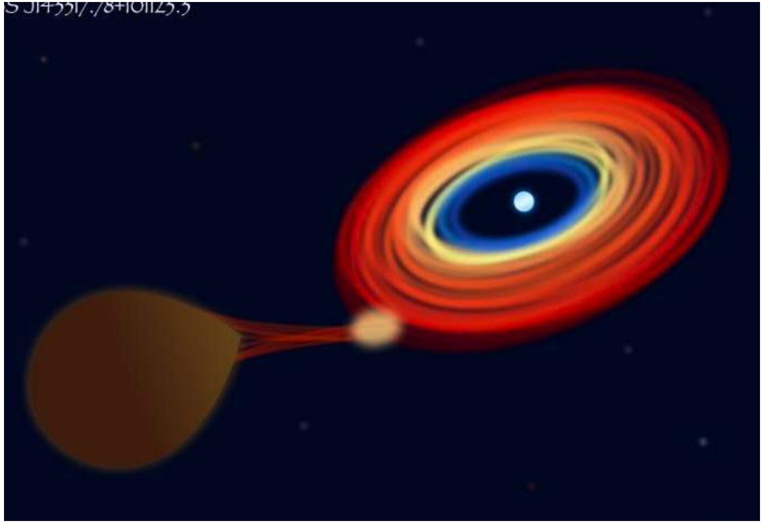
NERSC contact: Tuomas Koskela

NESAP

Synthetic Spectra of Astrophysical Objects

Studying three of the most exciting areas of current astrophysics: supernovae, extrasolar planets, and active galactic nuclei.

Modeling radiative transfer in atmospheres of very low mass stars and giant planets to understand observations.



Nature paper on discovery of an irradiated brown-dwarf companion to an accreting white dwarf.

PI: Eddie Baron, U. Oklahoma

NERSC Hours on KNL: 55 million

Max Concurrency: 8,064 nodes (548K cores)

Code: phoenix

NERSC contact: Brian Friesen

NESAP

Domain Wall Fermions and Highly Improved Staggered Quarks for Lattice QCD

Compute predictions of the standard model for kaon decays and mixings, quantities that offer highly-visible, exciting prospects for the discovery of physics beyond the standard model of particle physics.

PI: Norman Christ, Columbia University

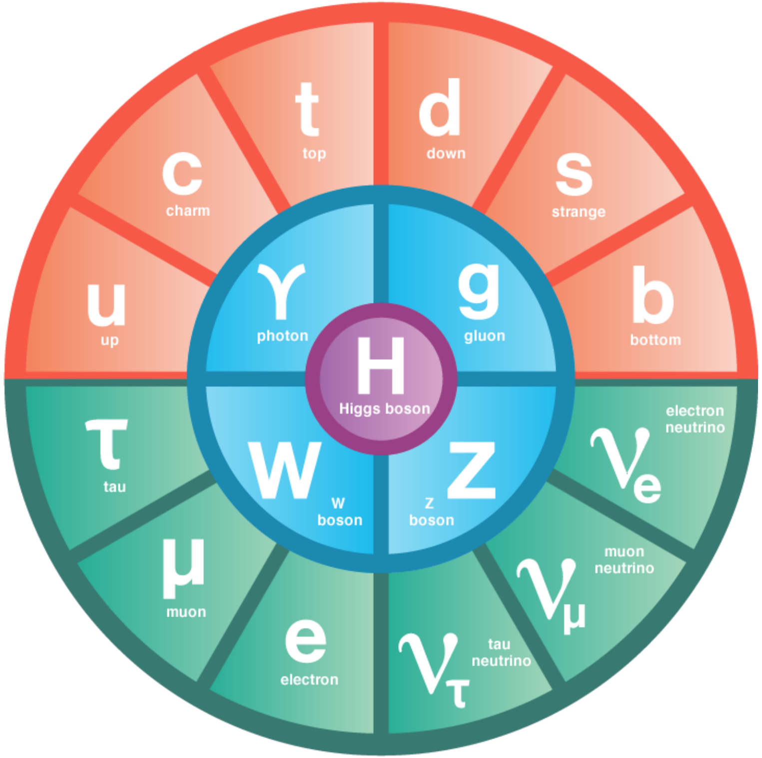
NERSC Hours on KNL: 430 million

Max Concurrency: 2,449 nodes (166K cores)

Code: DWF Inverter

NERSC contact: Woo-Sun Yang

NESAP



A Modern Computational Framework for the Nonlinear Seismic Analysis of Nuclear Facilities and Systems

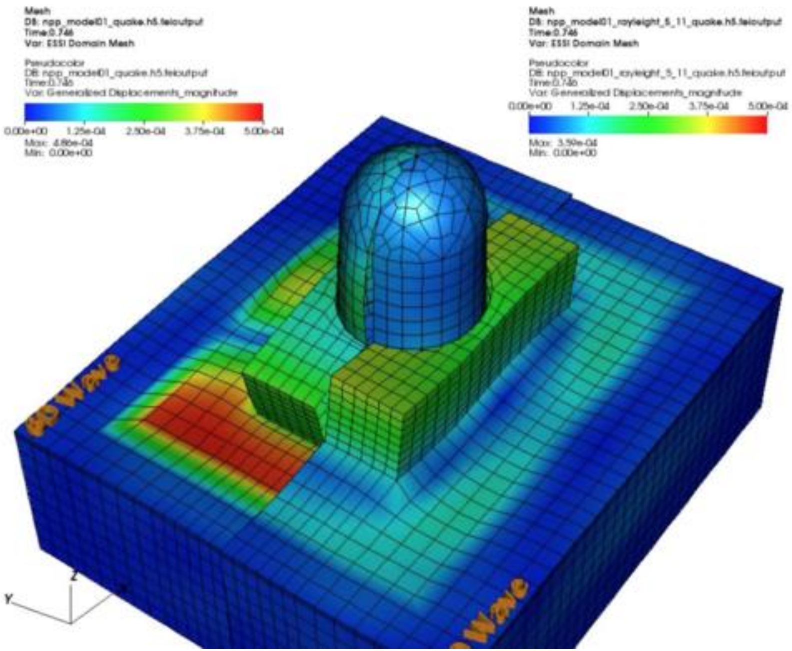
This project is focused on the development of advanced nonlinear modeling and simulation for seismic analysis of nuclear facilities. The ESSI nonlinear finite element program is being extended to included nonlinear structural elements necessary for a fully coupled nonlinear analysis of soil-structure systems.

PI: David Mccallen, Berkeley Lab

NERSC Hours on KNL: 8 million

Max Concurrency: 4,096mnodes (278K cores)

Code: SW4



Summary of NERSC Experiences

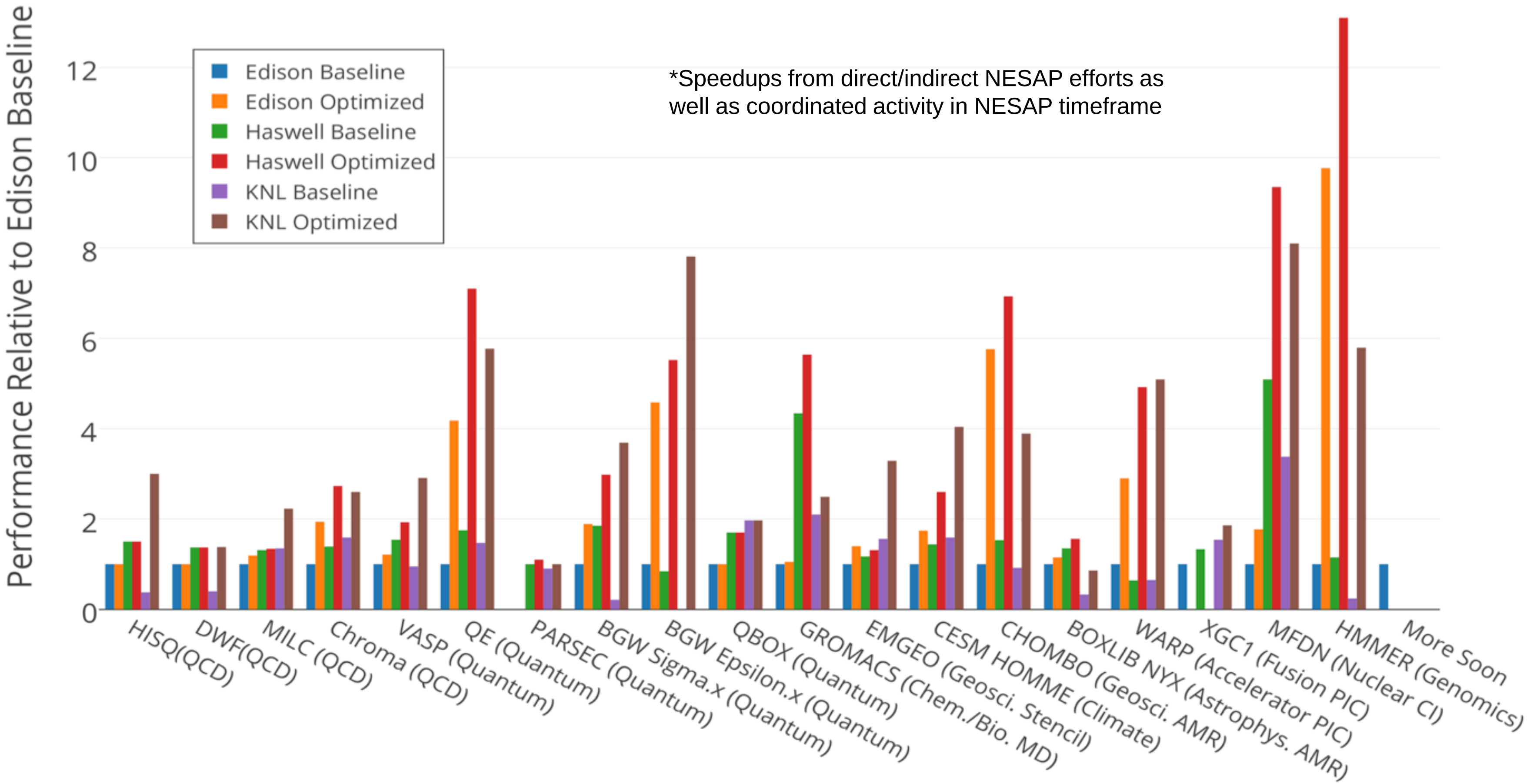
Cori with light-weight Intel Xeon Phi processors provides unprecedented capability for DOE Office of Science research

NESAP has enabled large percentage of NERSC workload to run efficiently on new class of manycore system

Lessons learned and knowledge gained are being communicated to and applied by NERSC community

Collaborations with application teams, vendors, and HPC community are necessary for success

NESAP Code Performance on KNL



*Speedups from direct/indirect NESAP efforts as well as coordinated activity in NESAP timeframe