NERSC

Cori KNL Update

IXPUG 2017

September 26, 2017

Richard Gerber

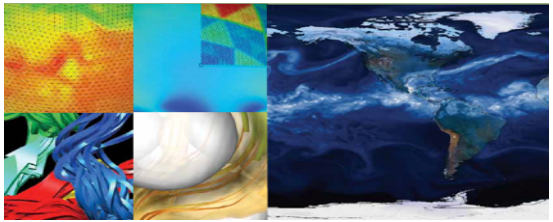
NERSC Senior Science Advisor

High Performance Computing Department Head

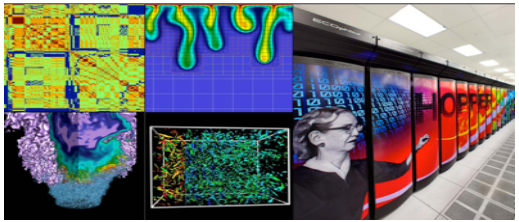
NERSC: The Mission HPC Facility for DOE Office of Science

U.S. DEPARTMENT OF ENERGY | Office of Science

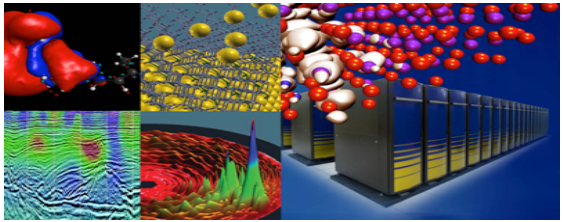
Largest funder of physical science research in the U.S.



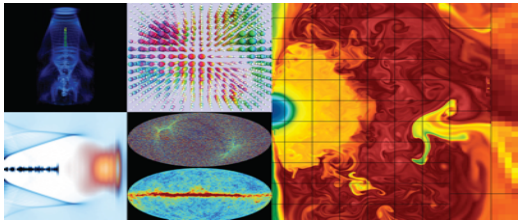
Bio Energy, Environment



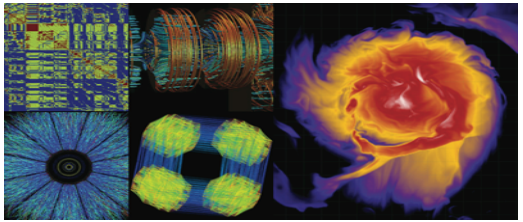
Computing



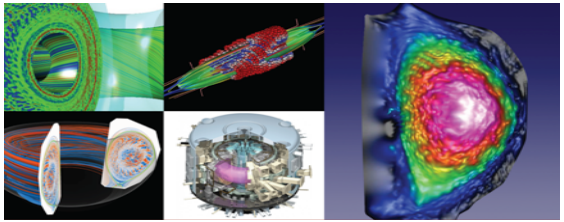
Materials, Chemistry, Geophysics



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 with 9,600 KNL nodes

68 core KNL / 272 HW threads

96 GB DRAM / 16 GB MCDRAM

Support the entire Office of Science research community

Transitioning workload to energy efficient architectures; on path to exascale

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 SLURM job scheduling

What is different about Cori for NERSC Users?

Edison (Cray XC w/ Intel Xeon Ivy-Bridge):

- 5000+ Nodes
- 12 Cores Per CPU
- 24 HW Threads Per CPU
- 2.4 GHz
- 8 DP Operations per Cycle
- 64 GB DDR Memory (2.6 GB/core)
- ~100 GB/s Memory BW
- 256b vector units
- 30 MB L3 cache per socket (12 cores)

Cori (Cray XC w/ Intel Xeon Phi KNL):

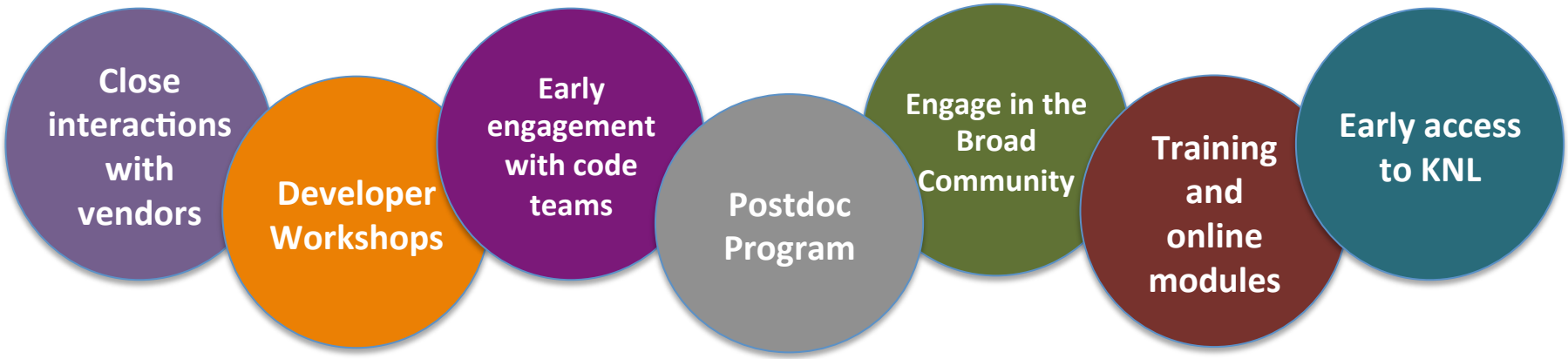
- 9600+ Nodes
- 68 Physical Cores Per CPU
- 272 HW Threads Per CPU
- 1.4 GHz
- 32 DP Operations per Cycle
- 16 GB of Fast Memory (0.24 GB/core)
  96GB of DDR Memory (1.4 GB/core)
  MCDRAM Has ~450 GB/s Memory BW
- No L3 cache
- 2 x 512b vector units

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

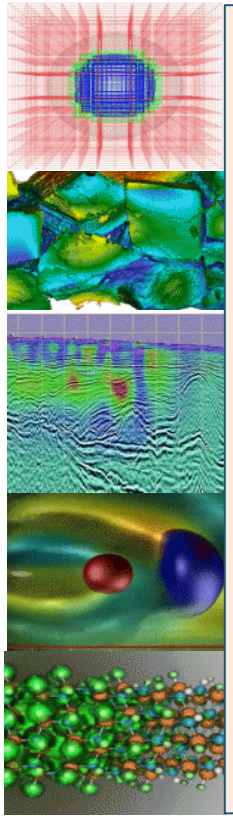
NERSC Exascale Scientific Application Program (NESAP)

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

Partner with ~20 application teams and apply lessons learned to broad user community – accounts for ~ 50% of hours used



NESAP Codes



Advanced Scientific Computing Research

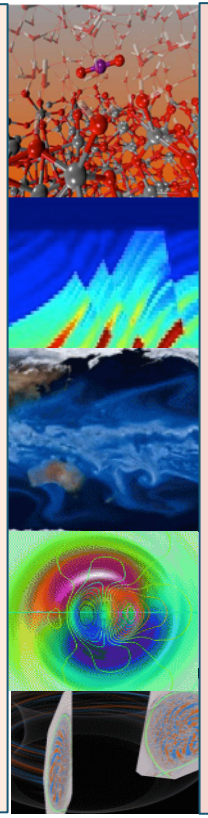
Almgren (LBNL) **BoxLib AMR**
Trebotich (LBNL) **Chombo-crunch**

High Energy Physics

Vay (LBNL) **WARP & IMPACT**
Toussaint(Arizona) **MILC**
Habib (ANL) **HACC**

Nuclear Physics

Maris (Iowa St.) **MFDn**
Joo (JLAB) **Chroma**
Christ/Karsch
(Columbia/BNL) **DWF/HISQ**



Basic Energy Sciences

Kent (ORNL) **Quantum Espresso**
Deslippe (NERSC) **BerkeleyGW**
Chelikowsky (UT) **PARSEC**
Bylaska (PNNL) **NWChem**
Newman (LBNL) **EMGeo**

Biological and Env Research

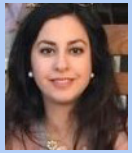
Smith (ORNL) **Gromacs**
Yelick (LBNL) **Meraculous**
Ringler (LANL) **MPAS-O**
Johansen (LBNL) **ACME**
Dennis (NCAR) **CESM**

Fusion Energy Sciences

Jardin (PPPL) **M3D**
Chang (PPPL) **XGC1**

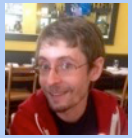
New Postdoc Program

Taylor Barnes
Quantum ESPRESSO



Zahra Ronaghi
Tomopy

Andrey Ovsyannikov
Chombo-Crunch



Bill Arndt
HIPMER/ HMMER/MPAS

Rahul Gayatri
SW4

Tuomas Koskela
XGC1

Kevin Gott
PARSEC

One Open Spot

NERSC Application Performance Group formed

New hire: Charlene Yang (Pawsey)

NESAP Staff Contributors

Katie Antypas

Jack Deslippe

Richard Gerber

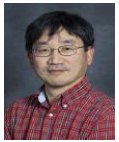
Nick Wright

Brandon Cook

Thorsten Kurth

Helen He

Stephen Leak



Woo-Sun Yang

Rebecca Hartman-Baker

Doug Doerfler

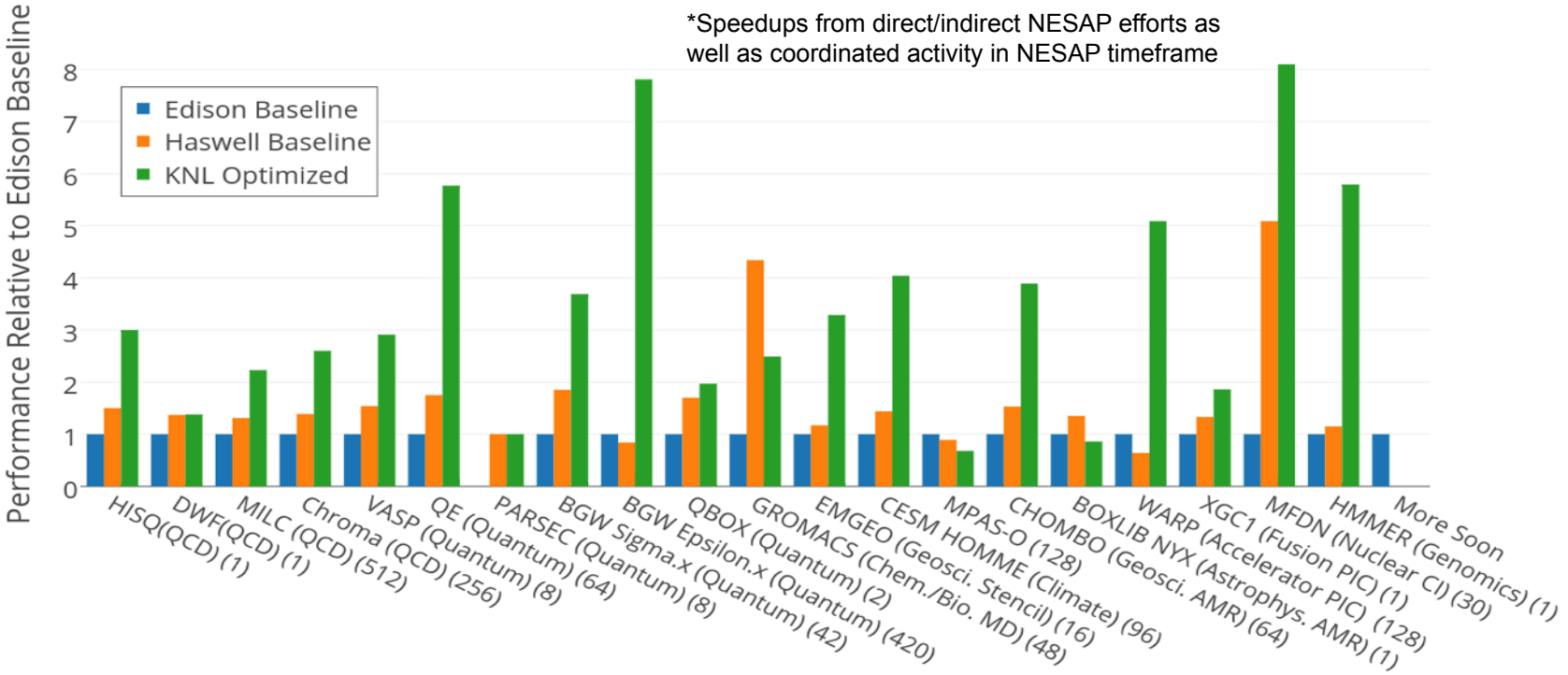
Zhengji Zhao

Brian Austin

Rollin Thomas

Brian Friesen
Former NESAP Postdoc

NESAP Code Performance on KNL



B: Baseline, original code
O: Optimized after NESAP work

H: Haswell dual core
I: Ivy Bridge dual core
K: Xeon Phi KNL

Ratio	Performance per node	Comment
HB/IB	1.6 X	Business as usual; not on path to exascale
KO/HB	2.5 X	NESAP + KNL benefit over Haswell no opt
KO/IB	3.6-4.0 X	Cori KNL optimized benefit over Ivy Bridge
HO/HB	2.3 X	NESAP code efforts only
KO/HO	1.1-1.2 X	Optimized KNL vs. optimized Haswell; on path to exascale
KB/HB	0.7 X	KNL vs. Haswell with no NESAP

KNL is 1.8X more energy efficient than Haswell (tentative results)

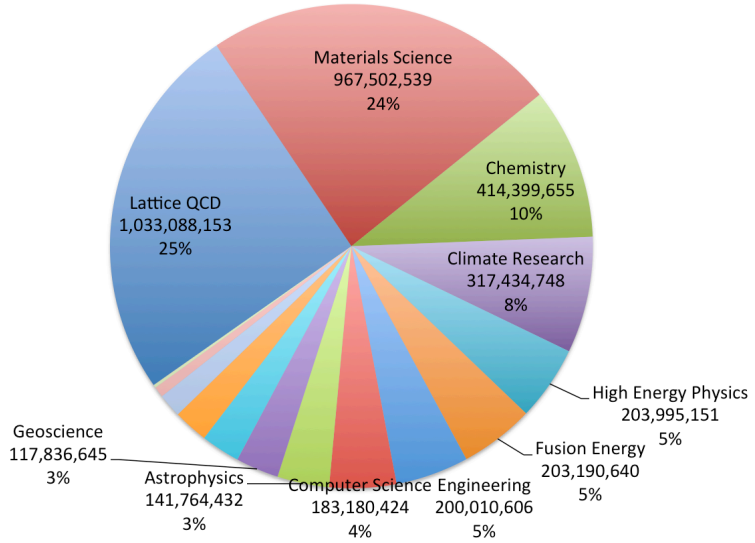
NESAP Has Enabled Broad Adoption of KNL

NERSC Exascale Science Application Program (NESAP) codes are running on KNL at about 3.5X their pre-NESAP performance on Edison per node.

138 projects have used > 1 M NERSC Hours on KNL

32% of hours used by jobs using > 1,024 nodes (69K cores)

NERSC supported 6 Gordon Bell submissions using Cori KNL



Cori provides a large increase in NERSC Hours available to Office of Science researchers at NERSC (3X+ in 2018 over 2016)

NESAP Has Enabled Broad Adoption of KNL

NERSC Exascale S
(NESAP) codes ar
their pre-NESAP p

138 projects have
KNL

32% of hours use
(69K cores)

NERSC supported
using Cori KNL

IXPUG has been a key resource for NESAP and NERSC

NERSC Hours
earchers at NERSC

(3X+ in 2018 over 2016)

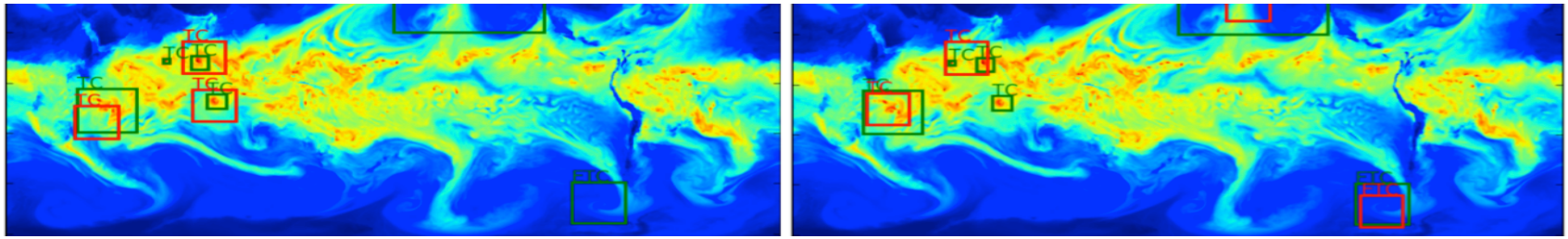
Deep Learning on Cori KNL

NERSC is actively exploring Deep Learning for Science

- Collaborating with leading vendors to optimize and deploy stack
- Collaborating with leading research institutions to develop methods
- Drive real science use cases

Deep Learning at 15 PF on NERSC Cori (Cray + Intel KNL)

- Trained in 10s of minutes on 10 terabyte datasets, millions of Images
- 9600 nodes, optimized on KNL with IntelCaffe and MKL (NERSC / Intel collaboration)
- Synch + Asynch parameter update strategy for multi-node scaling (NERSC / Stanford)



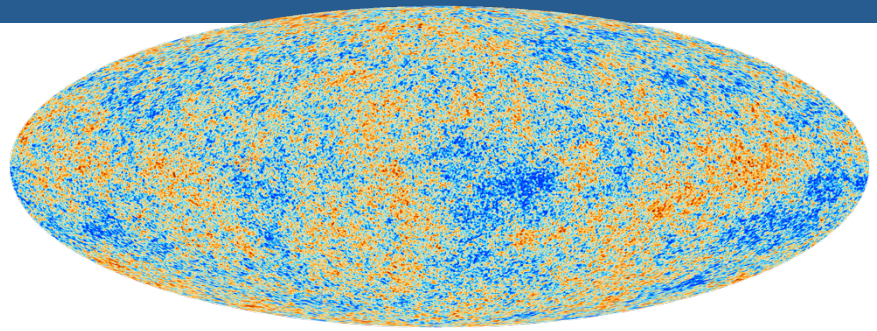
Identified extreme climate events using supervised (left) and semisupervised (right) deep learning. Green = ground truth, Red = predictions (confidence > 0.8). [NIPS 2017]

Scaling to 658K Cori Cores for CMB Analysis

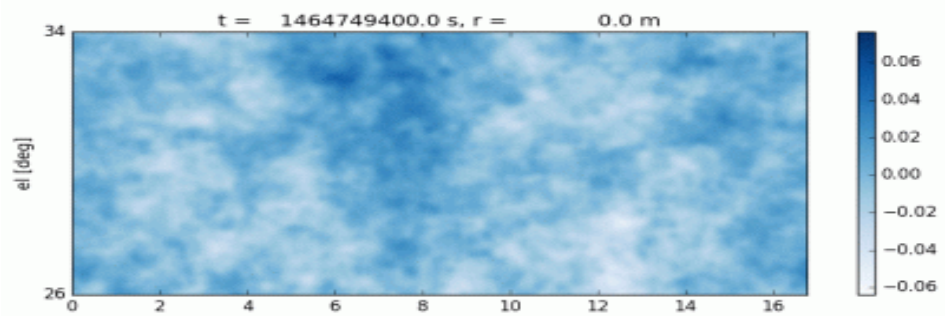
Berkeley Lab CS Computational Cosmology Center, NERSC, Intel, Cray collaborate via NESAP to scale key CMB software to full Cori system.

The TOAST (Time Ordered Astrophysics Scalable Tools) data simulation and reduction framework achieved a critical project milestone for upcoming experiments.

The ground-based CMB-S4 project will gather 35X more data than the Planck satellite did and will require TOAST's enhanced capabilities on Cori.



NERSC was used to reduce and interpret Planck results to create this map of the CMB



TOAST has incorporated modules to account for atmospheric effects

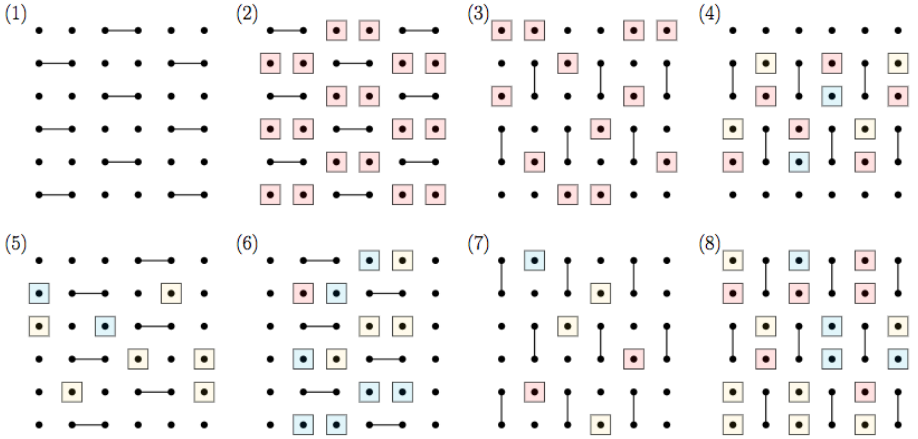
A Record Quantum Circuit Simulation

Scientific Achievement

Researchers from ETH Zurich used 8,192 Cori KNL nodes to successfully simulate a 45-qubit (quantum bit) quantum circuit, the largest simulation of a quantum computer achieved to date.

Significance and Impact

The current consensus is that a quantum computer capable of handling 49 qubits will offer the computing power of the most powerful supercomputers in the world. This new simulation is an important step in achieving “quantum supremacy”— the point at which quantum computers finally become more powerful than ordinary computers.



Low-depth random quantum circuit proposed by Google to show quantum supremacy.

Thomas Häner, Damian S. Steiger, 0.5 Petabyte Simulation of a 45-Qubit Quantum Circuit, arXiv: 1704.01127 [quant-ph]

Next steps and open issues

- Transfer KNL optimization lessons learned to the rest of the NERSC community of 7,000 users and 800 projects
- Scaling, I/O, and runtime variability (believed to be network-related) are still issues for some codes
- NESAP exploring application portability, novel programming models and runtimes, and influencing standards bodies
- NERSC 9 scheduled to arrive in 2020 and an active procurement is underway