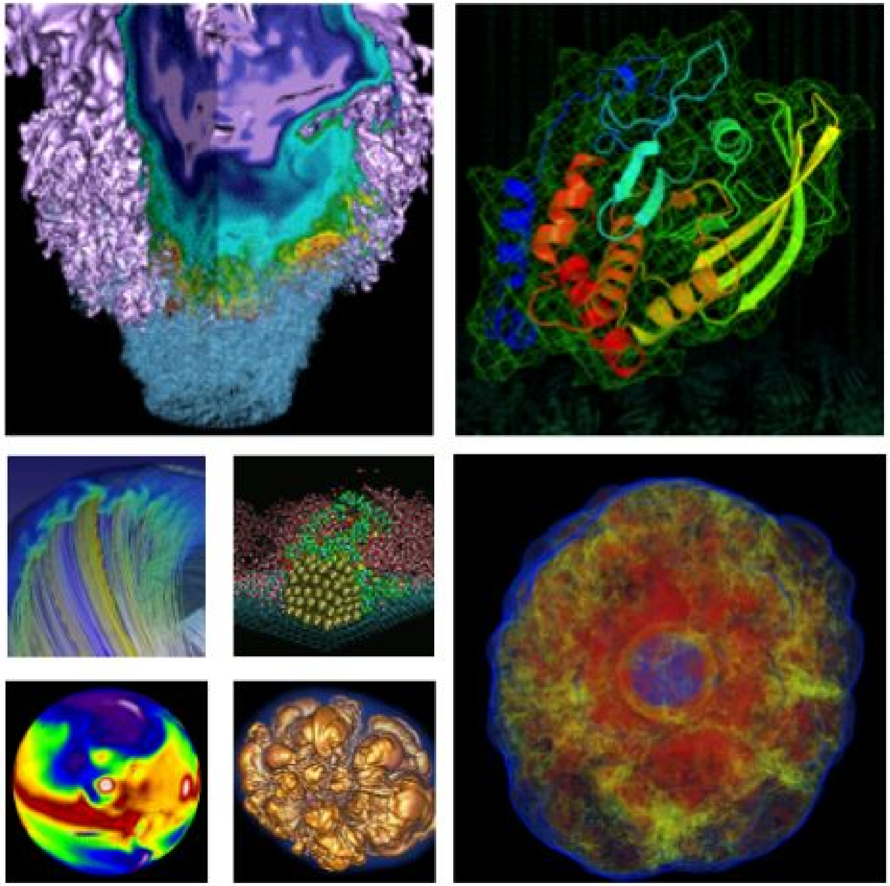
NUG October 2017 Meeting

- Today is the Great American Shakeout, an earthquake safety drill
- The drill started at 10:19 am
- If you're seeing this, NERSC is still in the drill and we'll start the webinar as soon as we get back
- Thanks for your patience!

NERSC Users Group Monthly Meeting



NERSC

October 19, 2017

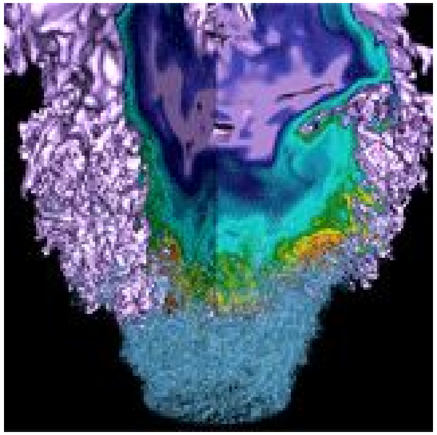
U.S. DEPARTMENT OF ENERGY

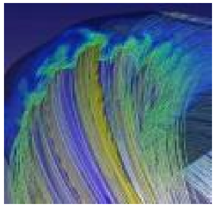
BERKELEY LAB
Lawrence Berkeley National Laboratory

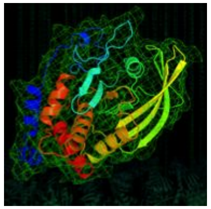
Agenda

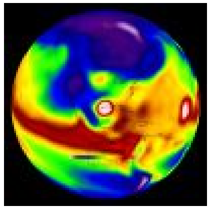
- Data Day/NUG 2017 recap
- Best Practices for I/O on KNL

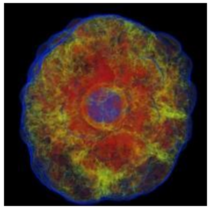
Data Day/NUG 2017 Recap

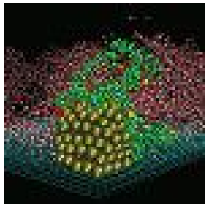












https://www.nersc.gov/users/NUG/annual-meetings/nersc-data-day-and-nug2017

Data Day 2017

- **NERSC's second annual Data Day was held on Tuesday 19th Sept**
  - data challenge/hackathon on Wednesday 20th.

- **13 speakers on a range of data-related topics from Deep Learning to data management**
- **Science areas included neuroscience, particle physics and LCLS.**
- **Over 80 attendees!**

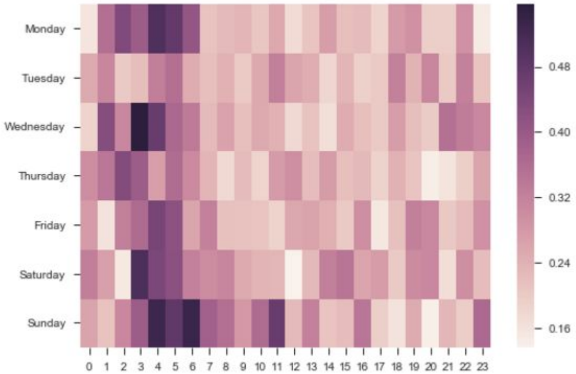
NERSC Supercomputing for Science and Energy
State-of-the art computing for the broad science community – over 7000 users, 700 applications
BERKELEY LAB

- **Recordings of all talks are on the website:**

https://www.nersc.gov/users/training/data-day/data-day-2017/

Data Competition

- **We challenged users to use NERSC machine learning tools to mine information from:**
  - SLURM job information
  - Astronomy dataset
- **Winning entries gave us real insight into our data!**
  - E.g. Lowest queue wait times if you submit on Saturday/Sunday night
- **Datasets and code from winning teams are up on the website**
  - Nice demonstration of using SciKitLearn and TensorFlow at NERSC



Fraction of low-wait jobs per hour per day (Juliette Ugirumurera and Liza Rebrova)

https://www.nersc.gov/users/NUG/annual-meetings/nersc-data-day-and-nug2017/data-competition/

NUG 2017 Day 1

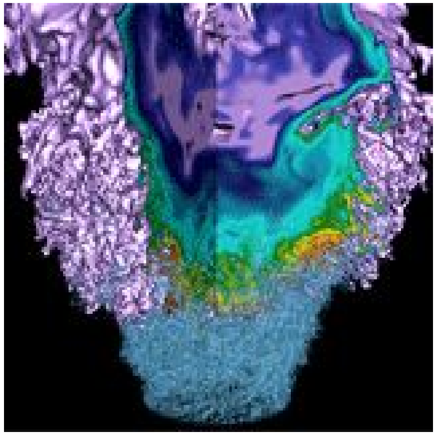
- **Morning: Data Competition, a combined event with Data Day 2017**
- **Afternoon: NERSC Status and Future Plans**
  - Views from DOE, NERSC Update, Innovations on Cori and Edison, NESAP2, Storage 2020, Big Data, Accounting and Security update, User Requirements and Survey
- **NUG 2017: 70 Attendees**
  - Slides available at https://www.nersc.gov/users/NUG/annual-meetings/nersc-data-day-and-nug2017/

NUG 2017 Day 2

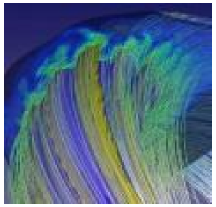
- **Featuring science and technology talks**
  - High Impact NERSC Sciences
  - NESAP
  - 5 large-scale Gordon Bell submissions
  - The best paper from IXPUG
  - A Keynote presentation from the ECP Director Doug Kothe on Exascale Science Applications
  - 4 NERSC HPC achievement awards (General and Early Career)
    - High Impact Science Achievement
    - Innovative Use of HPC

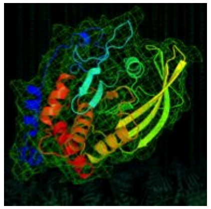
-- All 4 HPC award winners presented in person!
-- From Germany, ANL, LLNL, and LBNL
-- Details on HPC awards and winners are at: https://www.nersc.gov/news-publications/nersc-news/nersc-center-news/2017/nersc/

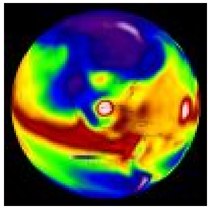
Best Practices for I/O on KNL

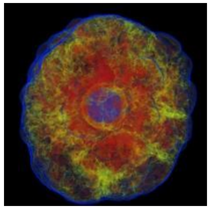


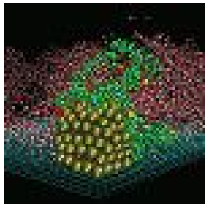
NERSC







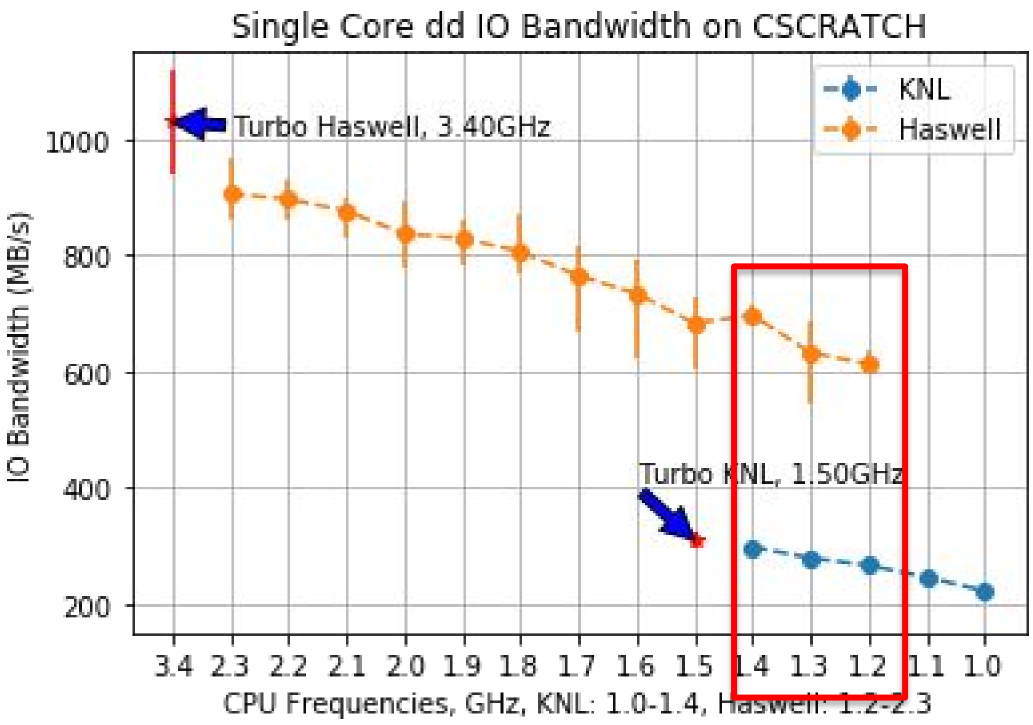




U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

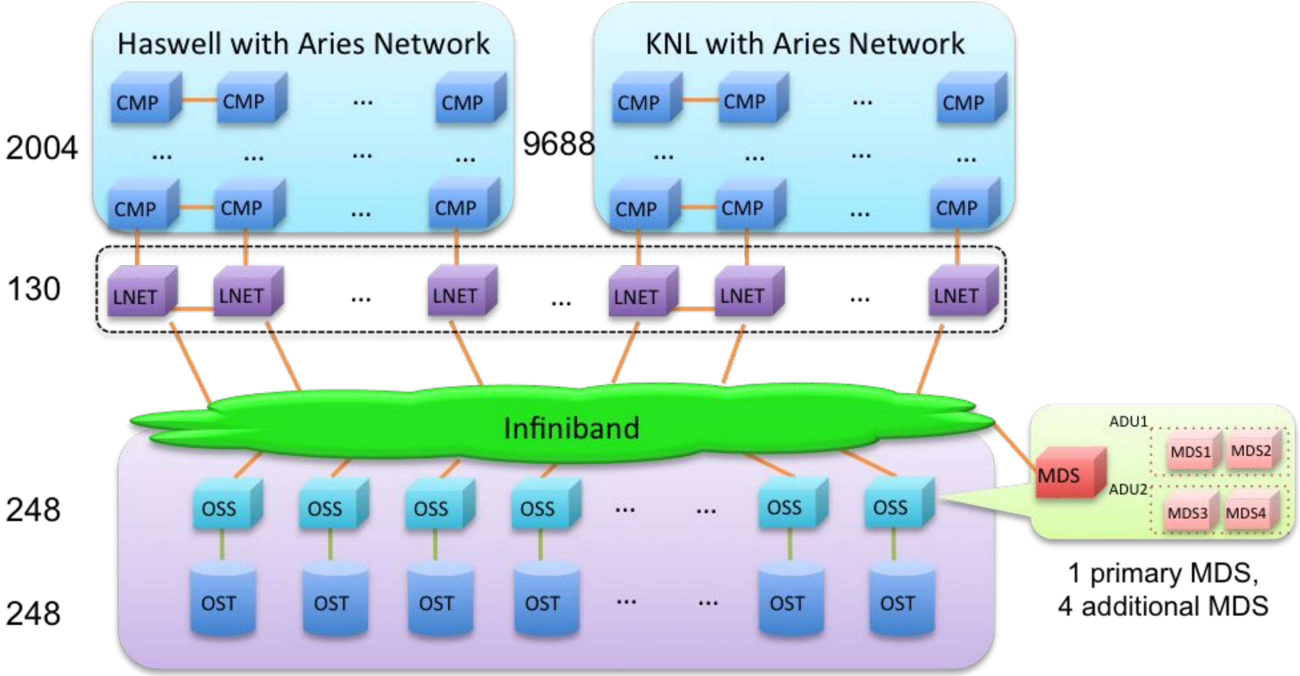
Single Stream IO Looks Bad…



- Bandwidth Ratio Haswell / KNL = **2.30** (at same CPU freq)

  = **3.46** (Turbo)

Parallel File System on Cori

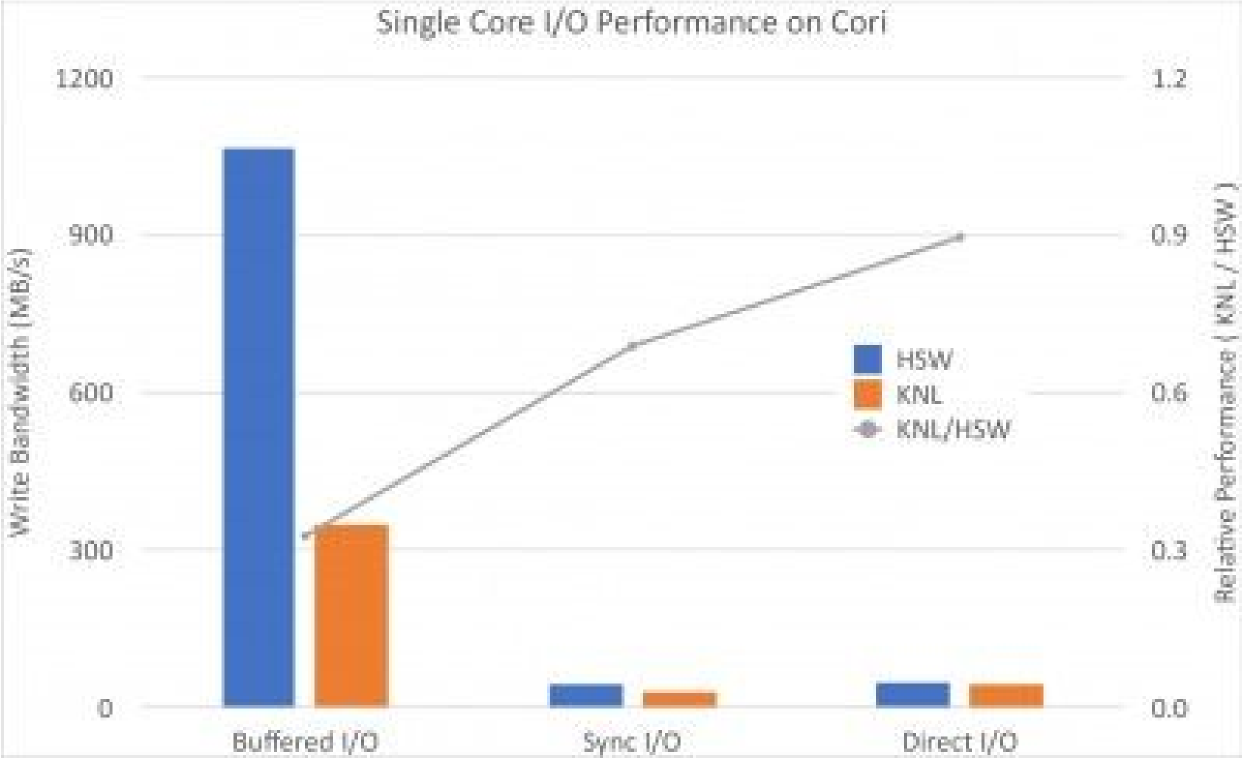


1 primary MDS,
4 additional MDS

Cori Haswell vs. KNL

	KNL	Haswell
CPU	1.4GHz	2.3GHz
Memory	96 G DDR4, 16G HBM	128 G DDR4
Cache(L1, L2, L3)	64K, 1M	64K, 256K, 40M
Node	68 core, single socket	32 core, two socket
Capacity	9688 nodes	2388 nodes

Cori Haswell IO vs. KNL IO



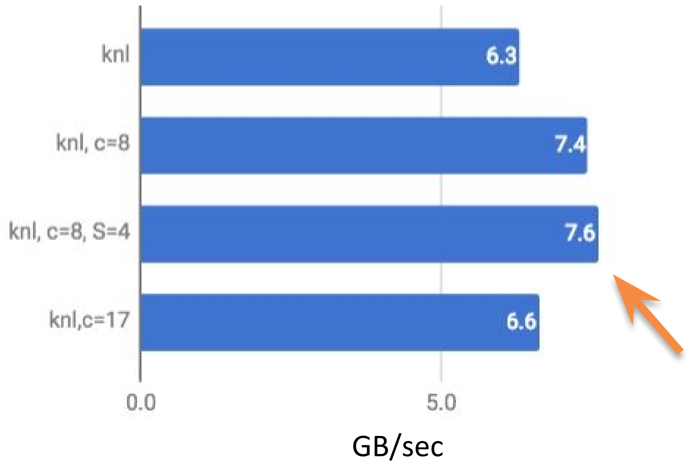
Major finding in the IO evaluation, CUG’17

Tip 1: Core Specialization

- Core Specialization

  #SBATCH -S 4

- Isolate system overhead to designated cores on a compute node.



Tip 2: Process Affinity

- Process Affinity, in case of node not fully packed, e.g., 4 MPI tasks

  use **--cpu_bind=cores** along with **–c**

  e.g., **srun -n 4 -c 64 --cpu_bind=cores**

- Otherwise, the processes will go to the same **core**

- Optimized WRF IO on KNL with 8X speedup

  300 seconds → 36 seconds (in reading initial input) -- John Michalakes, UCAR

Tip 3: Direct IO vs. Buffered IO

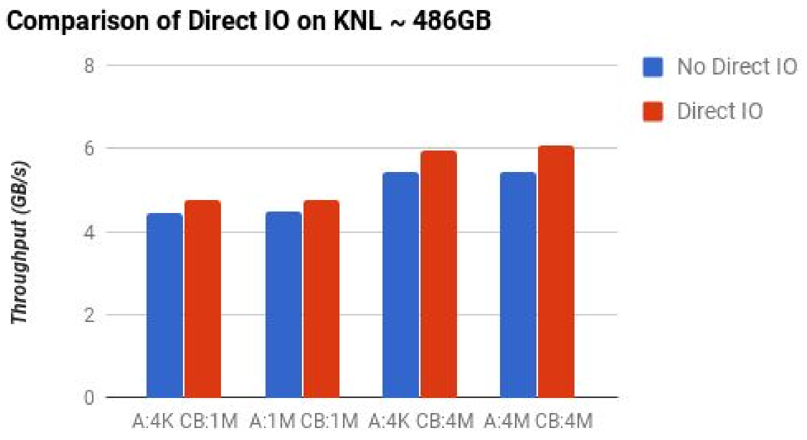
- KNL is close to Haswell with direct IO

  But Direct IO **largely slow** down your IO BW in most cases

- Page buffer benefits generally, e.g.**,** write, multi-read

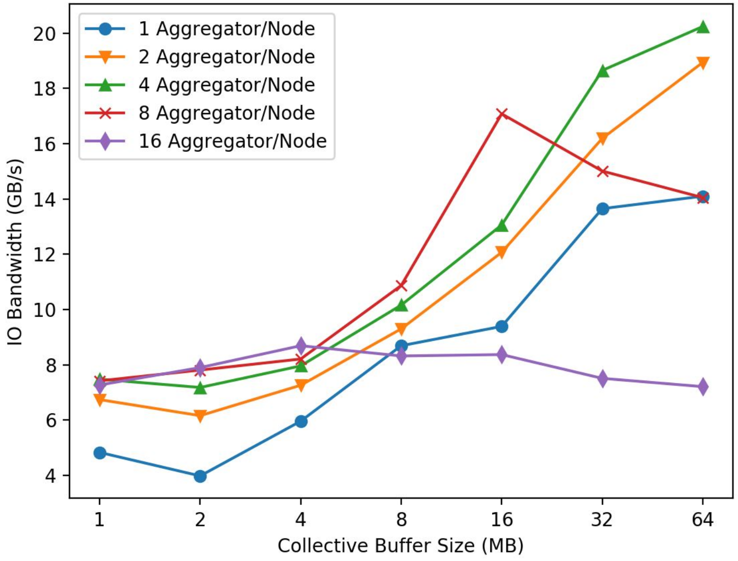
  User don’t need to do anything

- Direct IO can be better than buffered IO in **large one-time read**
  - POSIX: O_DIRECT in open()
  - IOR: -B
  - MPI: setenv MPIO_DIRECT_READ
  - HDF5: [H5Pset_fapl_direct()]()
    - 11% speedup

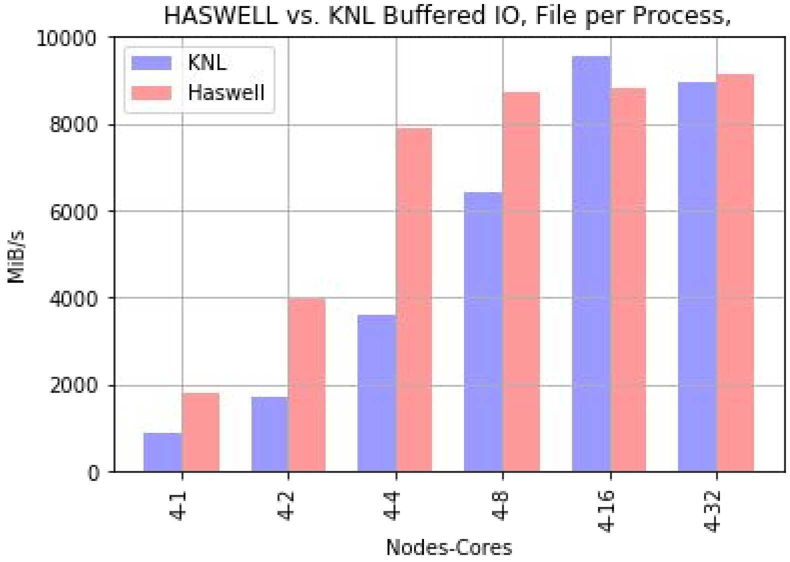


Tip 4: Collective Buffer

- KNL has larger inter-node latency than Haswell
- Increasing buffer size in MPIIO can improve IO BW



Tip 5.1: IO Parallelism - MPI



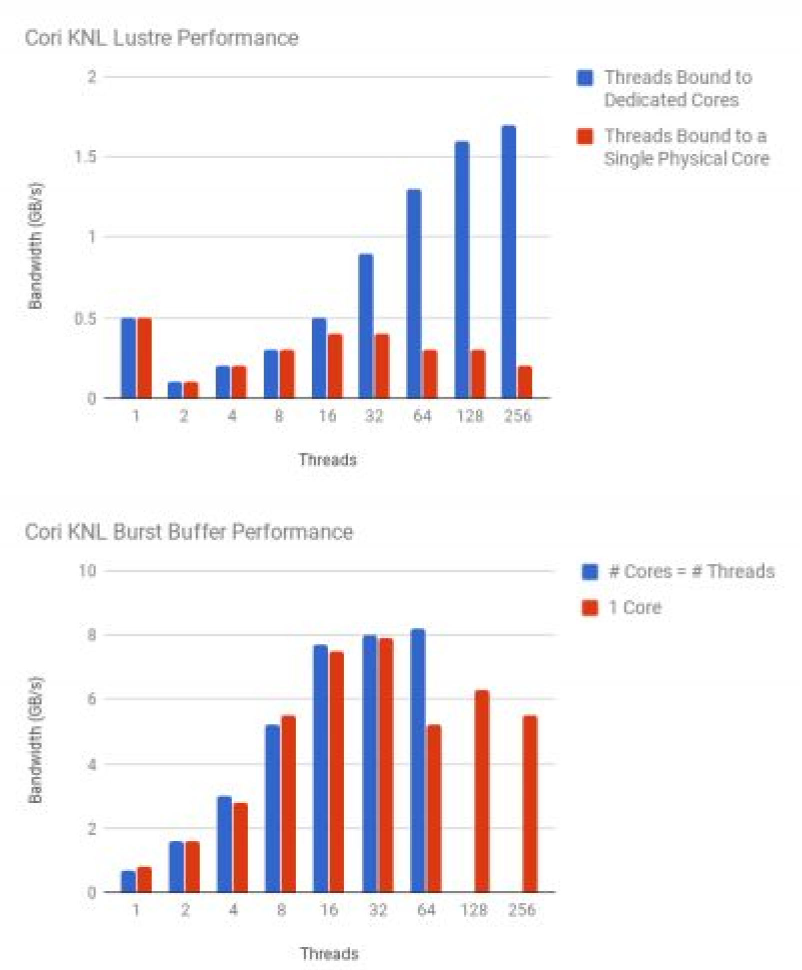
❖ KNL I/O performance can be markedly different from Haswell

❖ KNL requires multiple I/O streams to match Haswell performance.

❖ Initial results published at CUG’17.

J.L. Liu, Q. Koziol, H.J. Tang, F. Tessier, W. Bhimji, B. Cook, B. Austin, S. Byna, B. Thakur, G. Lockwood, J. Deslippe, Prabhat, Understanding the IO Performance Gap Between Cori KNL and Haswell, CUG’17

Tip 5.2: IO Parallelism - OpenMP



- Figure 1: 16 to 32 threads on a single core can saturate the bandwidth
- IO on burst buffer on a single KNL core is able to saturate the I/O bandwidth
- Benchmarks: Stanford Legion program

Tip 6: General Recommendations

- ❖ General Lustre striping and MPIIO optimization still apply to Cori KNL.

  lfs setstripe --size [stripe-size] --index [OST-start-index] --count [stripe-count] filename

https://www.nersc.gov/users/storage-and-file-systems/i-o-resources-for-scientific-applications/optimizing-io-performance-for-lustre/

Best Practices for KNL IO

- Core Specialization
- Process Affinity
- Direct IO vs. Buffered IO
- Collective Buffer
- IO Parallelism
- General Recommendations on Cori Lustre FS

https://www.nersc.gov/users/storage-and-file-systems/i-o-resources-for-scientific-applications/optimizing-io-on-cori-knl/

NERSC

U.S. DEPARTMENT OF ENERGY

Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory