



SC 19 Tutorial: Getting Started with Containers on HPC

Shane Canon¹, Sameer Shende², Carlos Eduardo Arango³, Andrew J. Younge⁴

¹ Lawrence Berkeley National Lab scanon@lbl.gov	² University of Oregon sameer@cs.uoregon.edu
³ Sylabs Inc eduardo@sylabs.io	⁴ Sandia National Labs ajyoung@sandia.gov



UNIVERSITY OF
OREGON



Sylabs.io



Sandia
National
Laboratories

exascaleproject.org



U.S. DEPARTMENT OF
ENERGY

Office of
Science

Outline

- 13:30 – 13:45 Introduction to Containers in HPC (Younge)
- 13:45 – 14:15 How to build your first Docker container (Canon)
- 14:15 – 14:45 How to deploy a container on a supercomputer (Canon)
- 14:45 – 15:00 Best Practices (Canon)
- 15:00 – 15:30 -- Break --
- 15:30 – 16:00 Running an HPC app on the E4S container (Shende)
- 16:00 - 16:30 How to build a Singularity container image (Arango)
- 16:30 - 16:50 Running Singularity on a supercomputer & adv features (Arango)
- **16:50 - 17:00 Success Stories & Summary (Canon)**

² Link: <https://tinyurl.com/yxbhpo35>

Success Stories & Summary

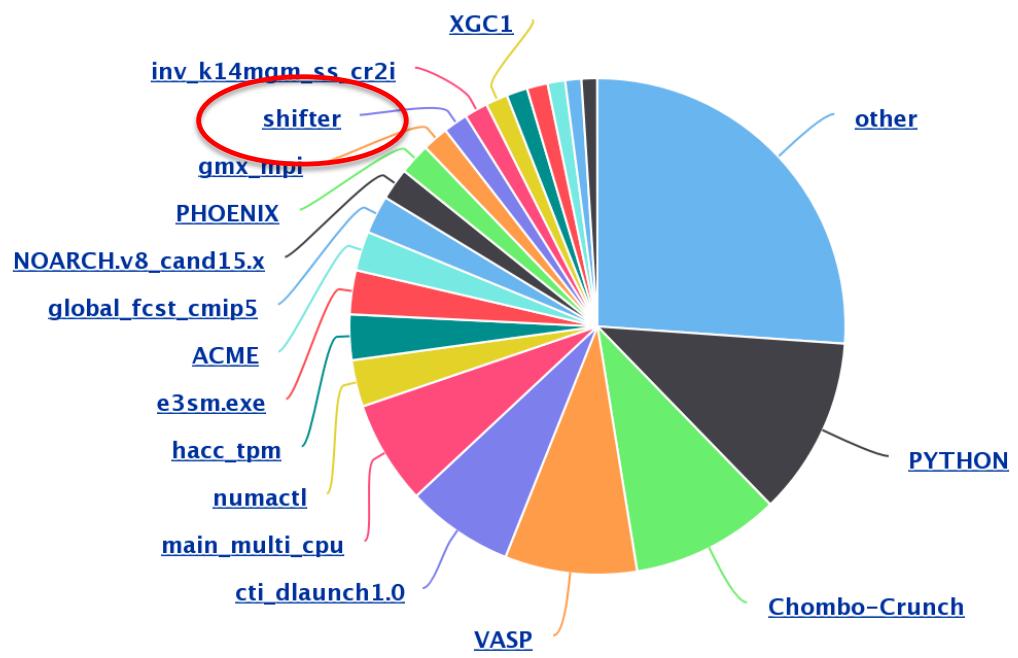


Stats on Shifter Use at NERSC

- 10M+ Image Lookups
- 1900+ Unique Image tags
- 700+ Unique Users
- Still a small fraction of NERSC overall use (~3%)

Cori Machine Hours Breakdown by Binary Names)

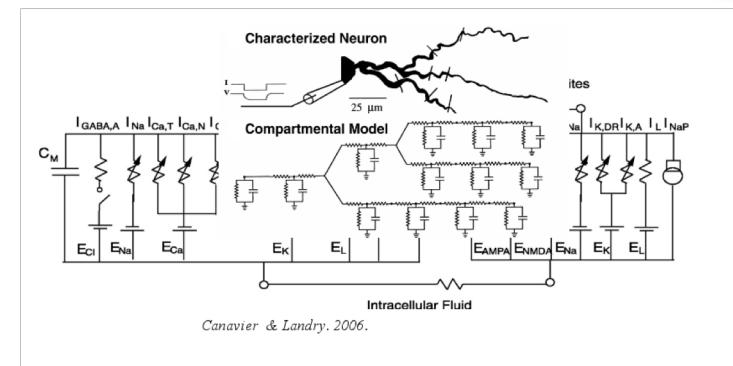
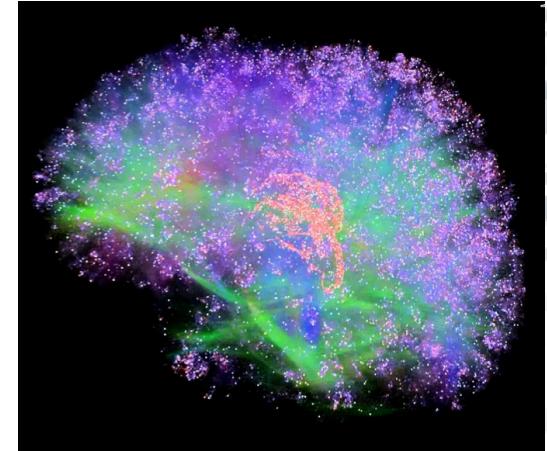
Processes as a percentage of 100% of total machine hours (5374916 hours).



Highcharts.com

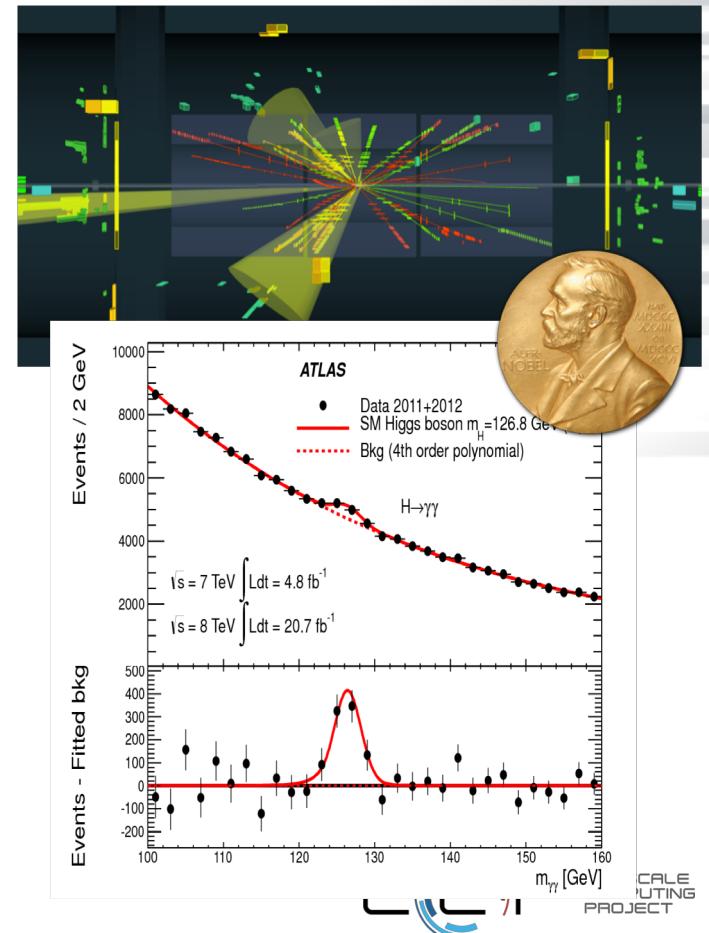
Modeling the Mind

- Neuron
 - Simulation program to model neuron response to stimuli
- Simulation: 3000 points by 4000 time steps for 400 neurons
- Why Containers?
 - Software developed in 1985, many older dependencies (32-bit)



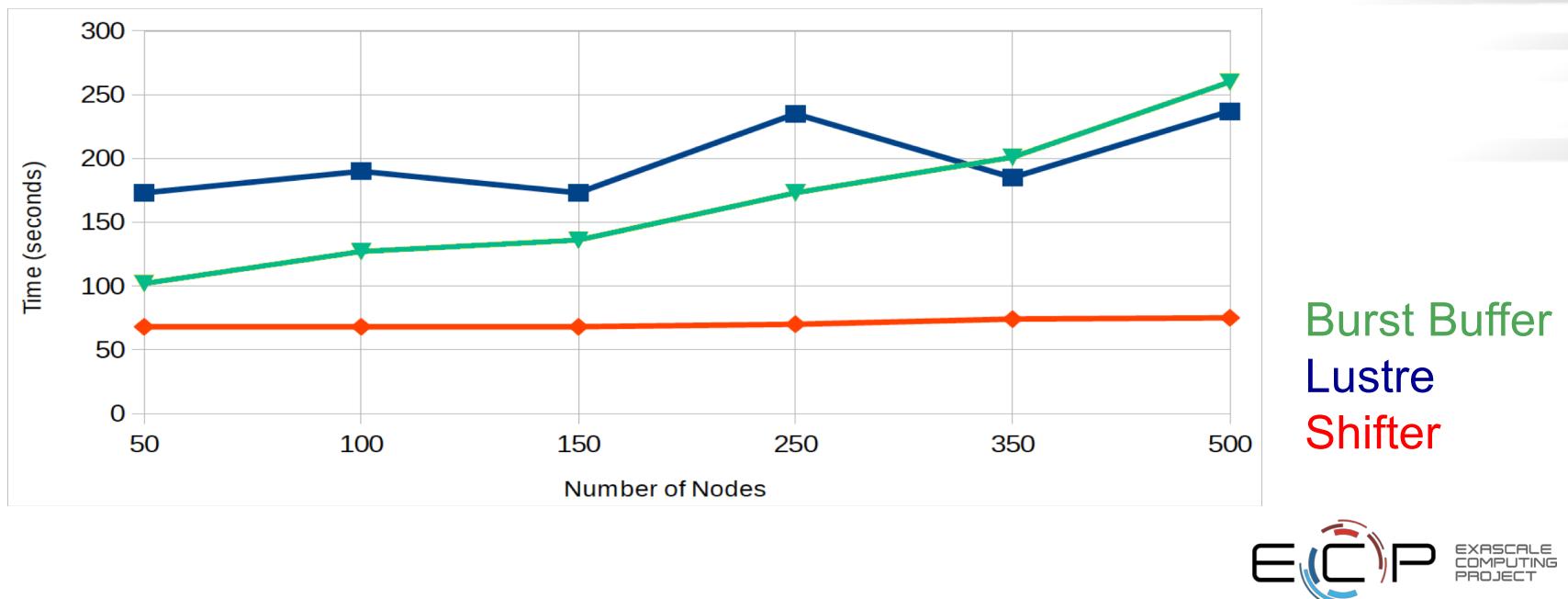
Probing the Fundamentals of Matter

- Large Hadron Collider (LHC)
 - 300 trillion proton-proton collisions and 30 PBs of data per year.
- Data analysis, simulation, multi-site data and computing pool
- Why Shifter?
 - Complicated software stack: Needs FUSE and elevated permissions to run
 - Integrated framework for running with images at all computing sites



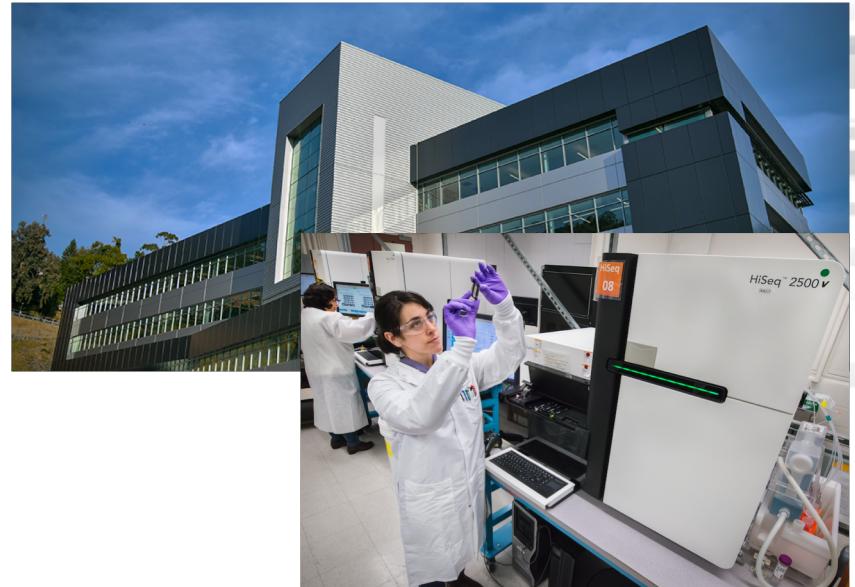
Creating a Monster

- The software stack is very big: 3.5 TB and 20M inodes
- Compress and de-duplicate with squashfs: 350 GB
- Start up time shows excellent scaling out to 500 nodes (16,000 cores)



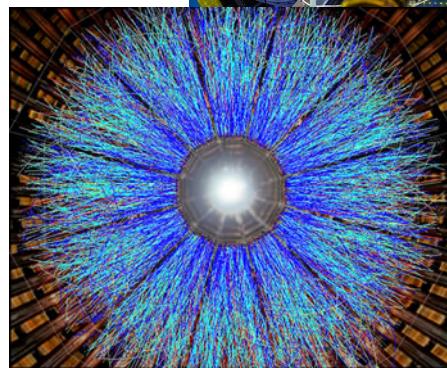
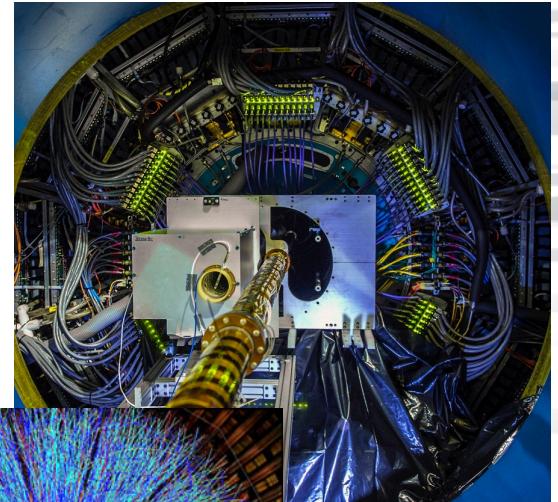
Joint Genome Institute

- Production sequencing center for DOE
- Sequencing microbial, microbial communities, plant and fungal organisms
- Why Containers?
 - Complex workflows utilizing dozens of tools for a single pipeline
 - Reproducibility and portability is critical
 - Container workflow tools (Cromwell) being used to encapsulate entire workflows



Probing The Nucleus

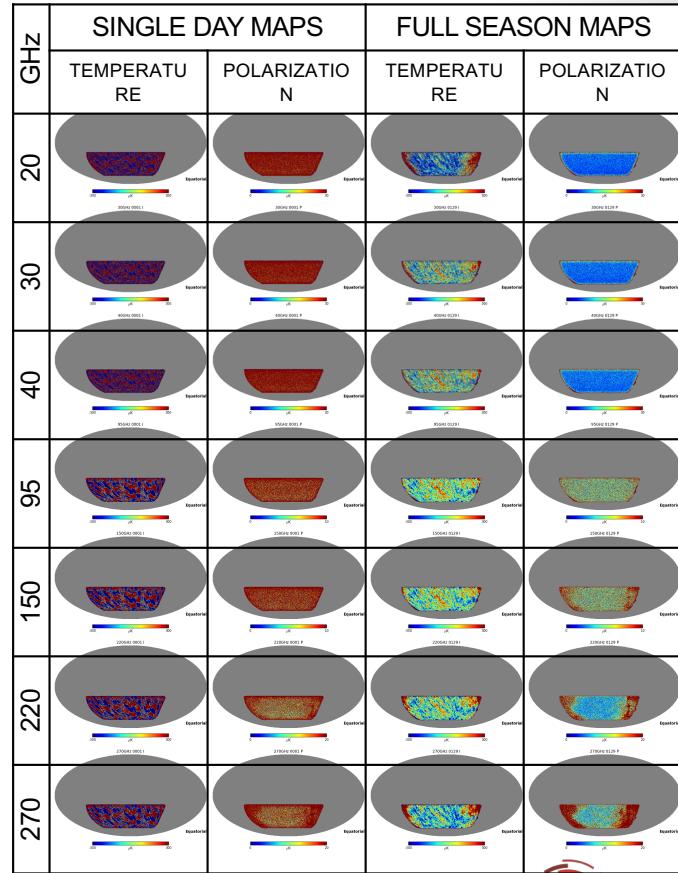
- STAR at Brookhaven, NY
 - smashing nuclei into each other to understand their component parts
- Data analysis and simulation
- Why Containers?
 - Difficult software dependencies (32-bit libraries)



Containers at scale

Measuring the Composition of the Universe

- **CMB – S4** - Ambitious collection of telescopes to measure the remnants of the Big Bang with unprecedented precision
- Simulated 50,000 instances of telescope using **600,000 cores** on Cori KNL nodes.
- Why Shifter and Containers
 - Python wrapped code needs to start at scale
 - Without Shifter application timed out during startup
- Uses Intel Compilers



Courtesy of Ted Kisner



Summary

- Thanks for your attendance and attention
- AWS Instances will be shutdown shortly
- NERSC Training accounts are active through the remainder of the day
- Please use the Google Doc or Slack to ask additional questions

Other SC activities

CANOPIE-HPC Workshop

- Time: Monday 2:30 PM – 5:00 PM
- Location: Rm 704-706

HPC Containers BOF

- Time: Wednesday 5:15 PM – 6:45 PM
- Location: Rm 401-402-403-404



Questions?

Next: learn how to work with your first container!



Spark

- “Big Data” high productivity analytics Framework
- Designed around commodity clusters (Ethernet network and local disk)
- Shifter image: lgerhardt/spark-1.6.0
- Uses per-Node write cache for spills and other temporary per-node file caches.
- Tested up to full scale of Cori Phase 1 (1600 nodes) with multiple Spark applications.

