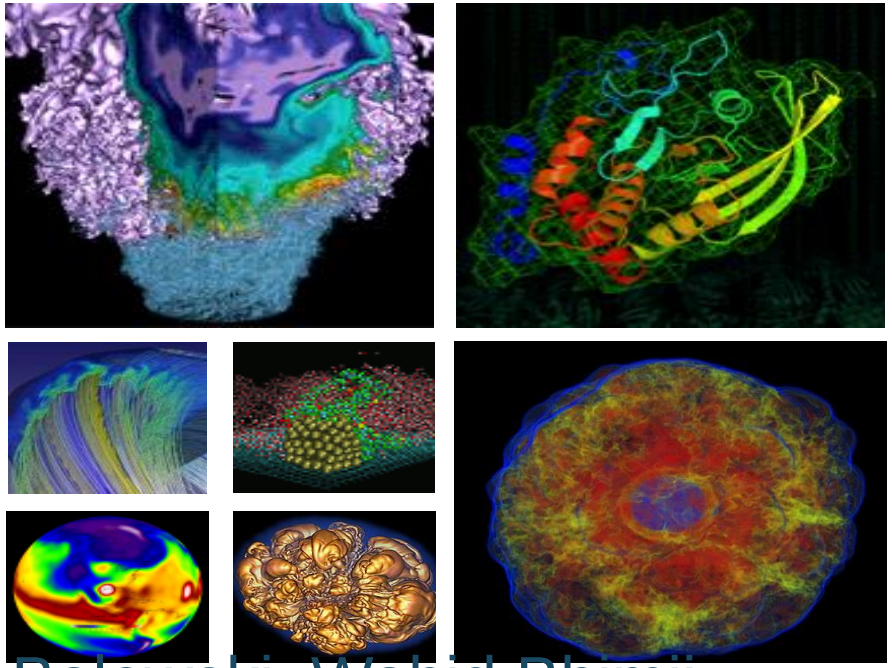
Enabling production HEP workflows on Supercomputers at NERSC

CHEP 2018
July 9th 2018



Jan Balewski, Wahid Bhimji, Shane Cannon, Lisa Gerhardt, Rei Lee, Mustafa Mustafa and others @NERSC Berkeley Lab (LBNL)

NERSC

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

Outline

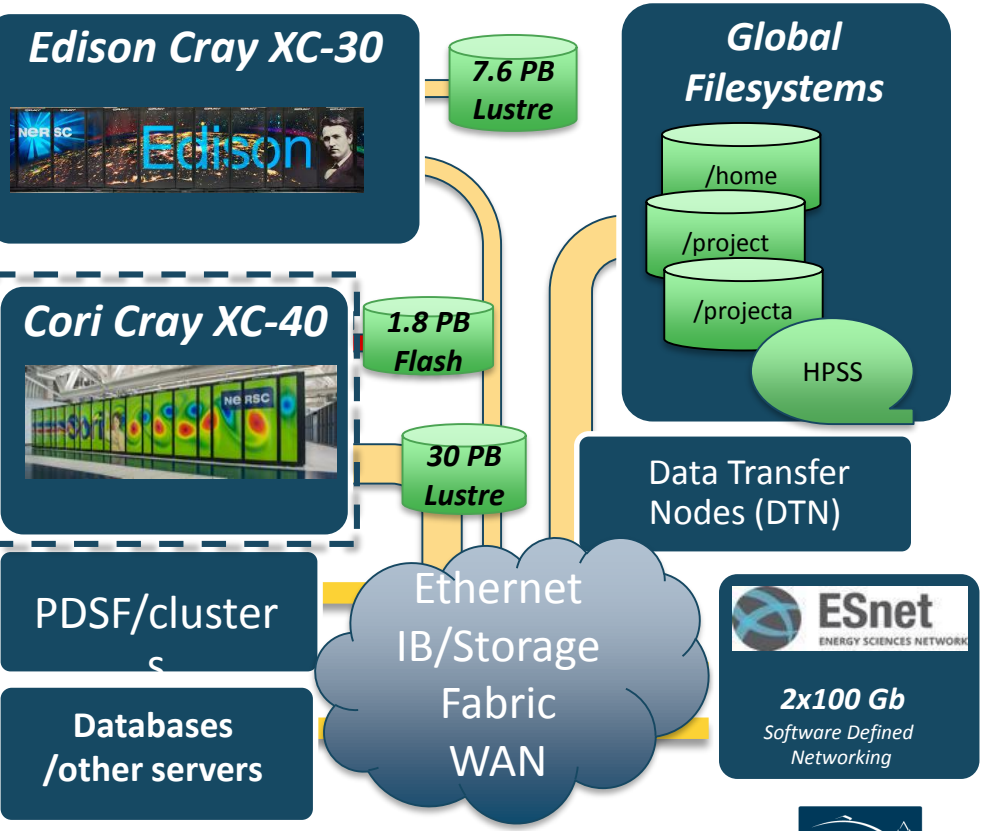
- **Introduction to NERSC**
- **HEP workflow components:**
  - Challenges on HPC; Approaches; NERSC technologies
- **Some recent Experimental-HEP@NERSC workflow highlights:**
  - Achievements; Approaches; Observations from user-support perspective
- **Some recent technology developments:**
  - DVFMS; Cori Networking; SPIN
- **Future directions for NERSC:**
  - NERSC-9; Storage; Workflows

NERSC

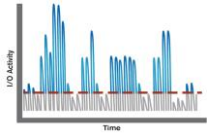
Mission HPC center for US Dept. of Energy Office of Science:
>7000 users; 100s of projects; diverse sciences

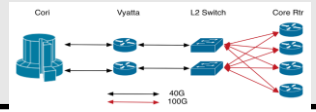
Cori: 31.4 PF Peak –#10 in Top500

- 2388 Haswell 32-core 2.3 GHz; 128 GB
- 9668 KNL XeonPhi 68-core 1.4 GHz
  4 hardware threads; AVX-512;
  16 GB MCDRAM, 96 GB DDR4
- Cray Aries high-speed “dragonfly” interconnect
- 28 PB Lustre FS: 700 GB/s peak
- 1.8 PB Flash Burst Buffer: 1.7 TB/s

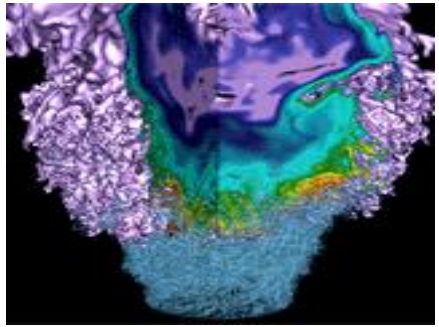


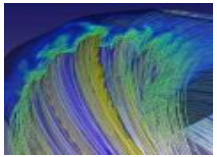
Workflow Component	Possible Issues	Approaches and Tech @ NERSC
Software: • Base OS • Experiment	Cray Linux No fuse for cvmfs Shared filesystems	Containers: Shifter CVMFS with Cray DFS -> 'DVFMS' Read-only /global/common filesystem
IO • Bulk data • Small files	IOPS and metadata on shared Lustre filesystem	Lustre DNE on Cori Burst Buffer Shifter Per-node-cache
Databases/ Services	Limited server capacity or access	Remote access; Read-only copy (shifter); On-site (SQL,MongoDB)/SPIN;
Workflow: • Job submission • Orchestration	Queue policies Server access	Scripts on Login/Workflow Nodes/SPIN SLURM (flexible queues) Grid services; 'Bosco (ssh)'
Data Transfer • Scheduled • In-Job	Compute nodes on separate High-Speed Network	Scheduled: Data Transfer Nodes (DTN) In Job: 'SDN'

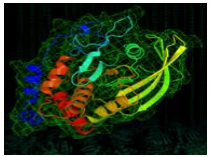


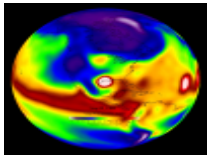


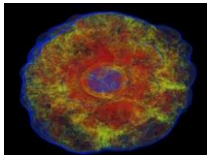
Some production runs from last year and observations

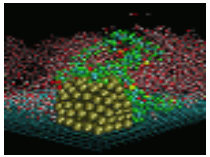








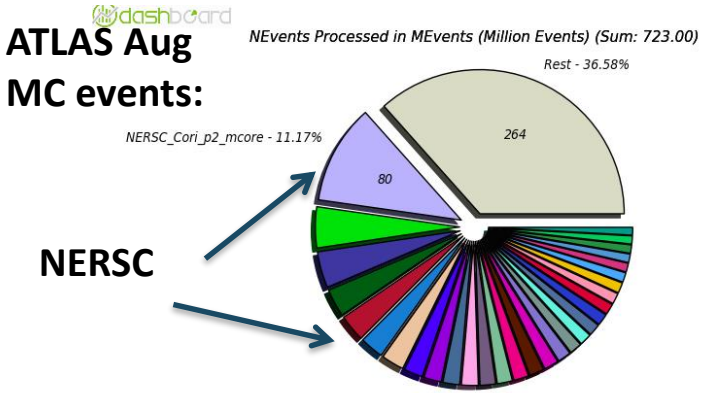




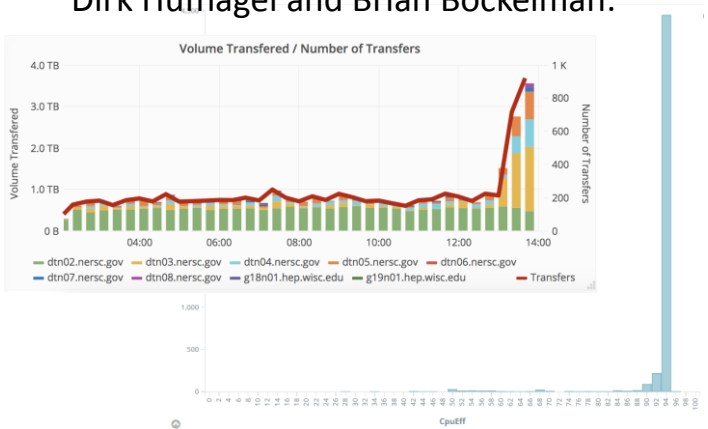
ATLAS / CMS: Production Integration

- **ATLAS: Mostly MC production**
  - ATLAS submission account was a top 3 NERSC "user" in 2017 and NERSC a top ATLAS MC producer
  - Pilots sub from login/workflow nodes (now using 'Harvester'); jobs with various size/ times avoid Qs
- **CMS: Many workflows**
  - Remote reading of pileup files from Fermilab: helped drive Cori node external networking – but still saw some residual connection timeouts
  - Copied pileup to NERSC (via DTNs/rucio). Local read has good CPU efficiency so running like that but still will explore remote read further
- **Both also using/stressing SPIN for frontier/squid servers and DVMFS and Shifter per-node-cache**

ATLAS Aug MC events:



Dirk Hufnagel and Brian Bockelman:

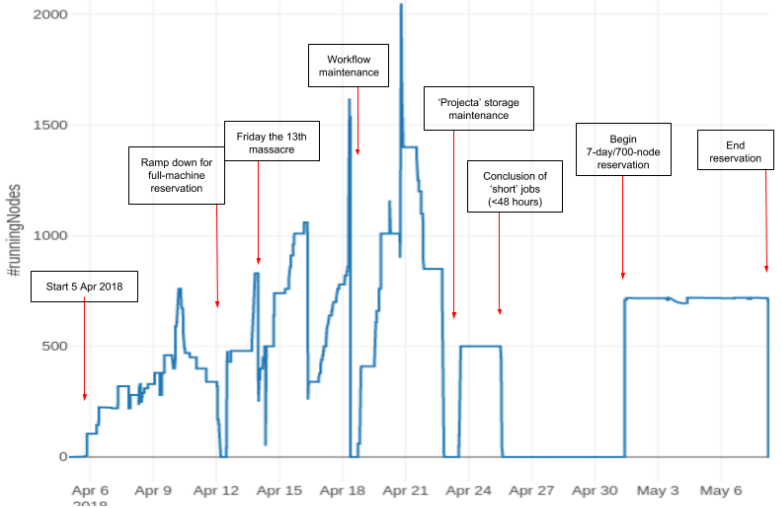


LSST-DESC/LZ: Data Challenges

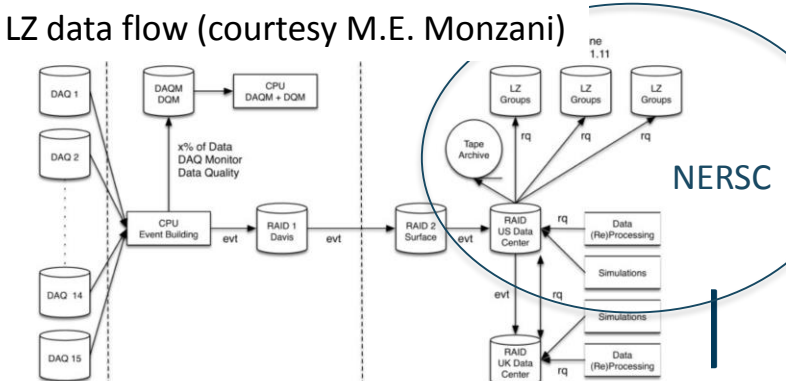
- **LSST-DESC data challenge (DC2) : phoSim image generation ('Raytrace')**
  - Uses 34 processes each of 8-threads on KNL
  - Issues include: 48hr pilots can't backfill – long queue wait times. Also have jobs > 48hr
  - Reservation allowed longer jobs to progress

- **LZ: Plan to run several workflows at NERSC**
  - Recent data challenge (MDC2) – 1M jobs
  - Memory capacity limited so Edison best 'value'
  - Using DVMFS with /project mount as backup
  - I/O to /project - issues at >1000 node scales

Plot from DC2 issue tracker – courtesy T.Glanzman



LZ data flow (courtesy M.E. Monzani)

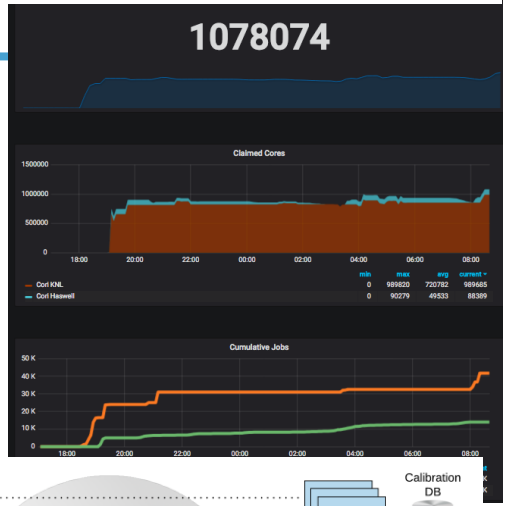


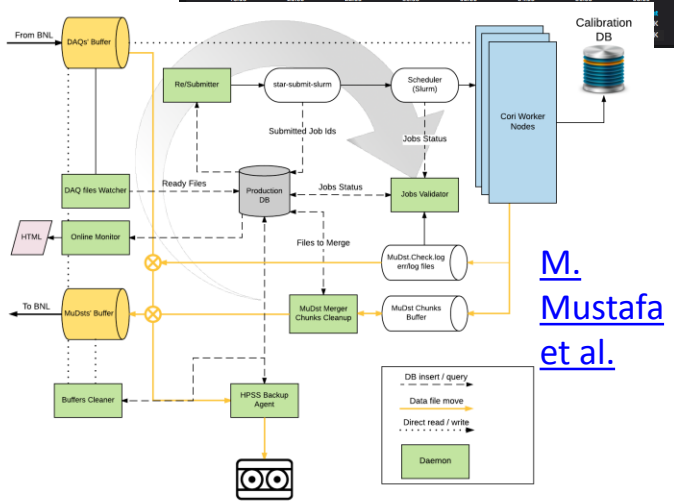
NoVA/STAR: Large-scale process...

- **NoVA: Multi-dimensional neutrino fits**
  - 1m cores in reservation across all Cori (hsw and knl)
  - 35M core-hours in 54 hr total
  - Fast turn around of processing (via reservations) for Neutrino18 Conf

- **STAR: Data reconstruction:**
  - Transfer from BNL via DTN
  - Very efficient stateless workflow:
    - >98% prod eff
    - Use local MongoDB
    - MySQL read-only DB in shifter per-node-cache

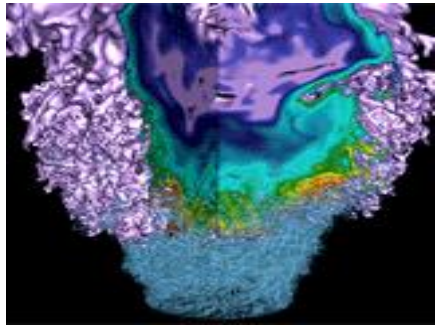
Nova HepCloud Monitor (courtesy Burt Holzman)

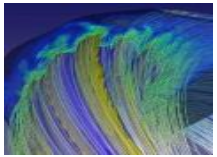


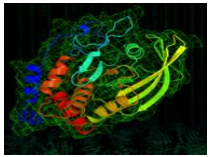


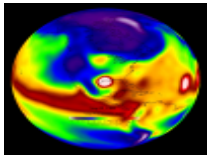
M. Mustafa et al.

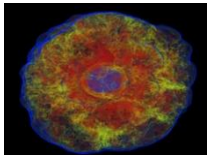
A couple of recent technology developments

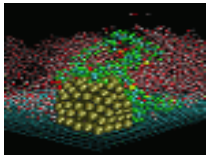








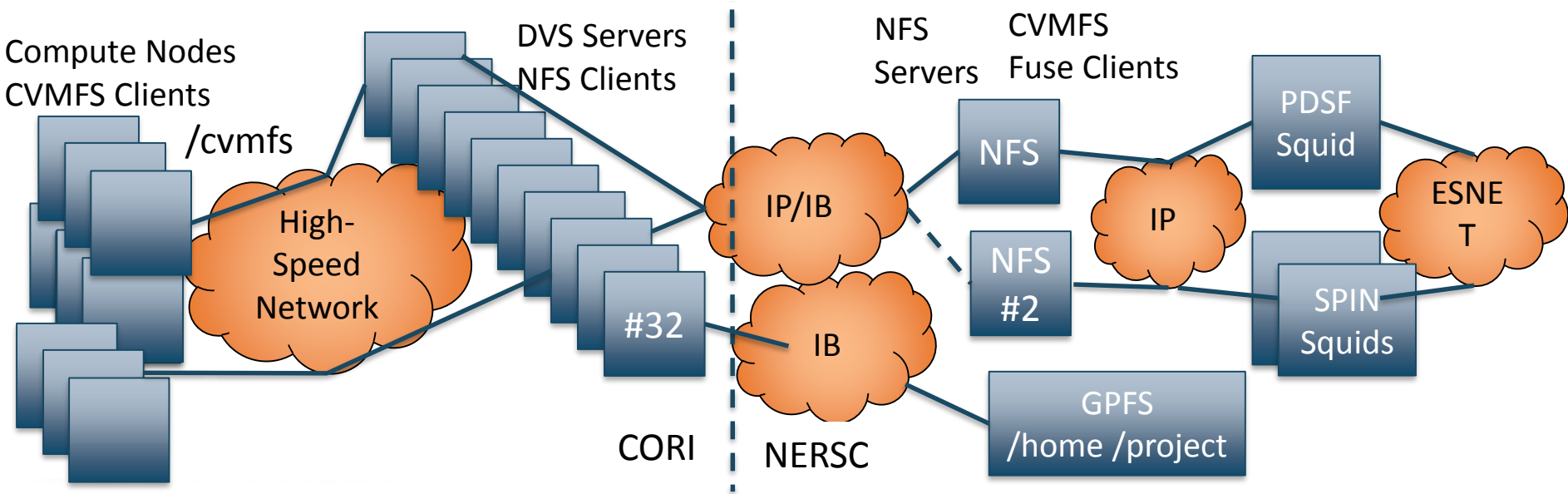




CVMFS -> DVMFS

Rei Lee, Lisa Gerhardt,
Shane Cannon …

- Restrictions with compute OS (FUSE etc.) has made providing /cvmfs at NERSC painful:
  - Can stuff into shifter containers – used in production by ATLAS/CMS
  - But large images; non-instant updates; adding other releases/repos not easy etc.
- Instead use Cray DVS (IO forwarder for non-lustre filesystems) to provide up-to-date CVMFS (over NFS) with caching at DVS nodes



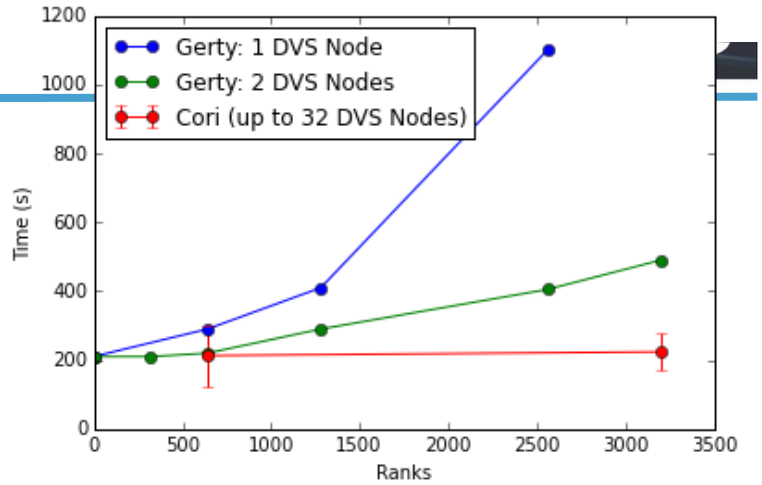
DVMFS

- **Startup time scales fine (with enough DVS Servers)**

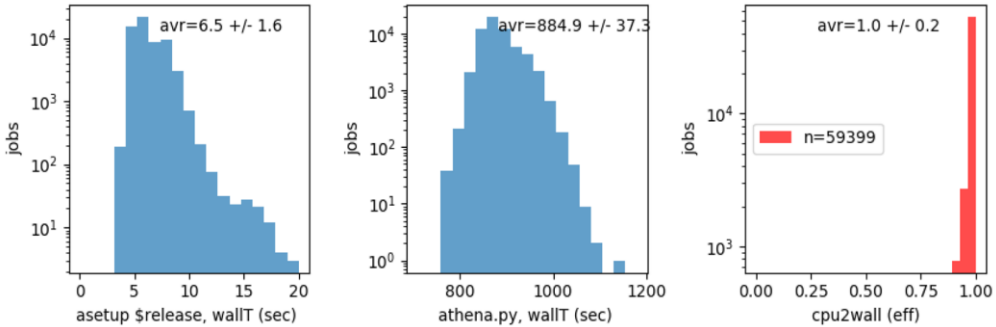
Many issues encountered on Cori:

- **Cray kernel bug / DVS patch**
- **Excessive boot time for mounts**
  - Use crossmnt of /cvmfs
- **Receive wrong file! (#1)**
  - 2 NFS servers have different inodes
- **Receive wrong file! (#2)**
  - Different repos reuse inodes and because of dvs and crossmnt can clash
  - Back to separate mounts
- **Now seems stable against errors**
  - 16 repos mounted.

Atlas simulation: Time to first event (test system (Gerty) and Cori:

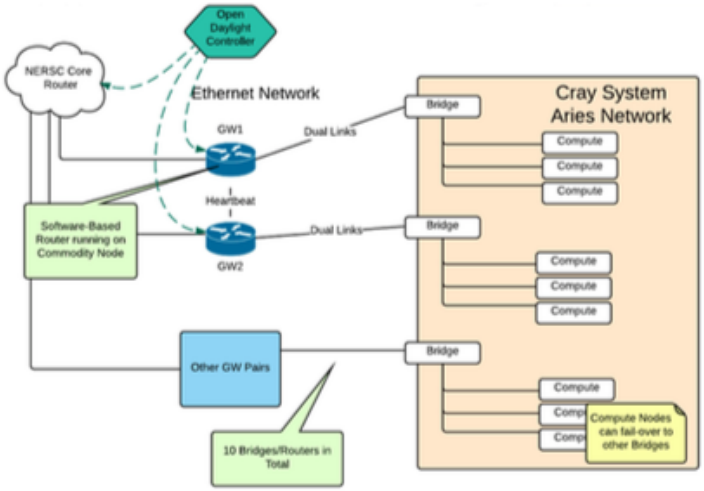


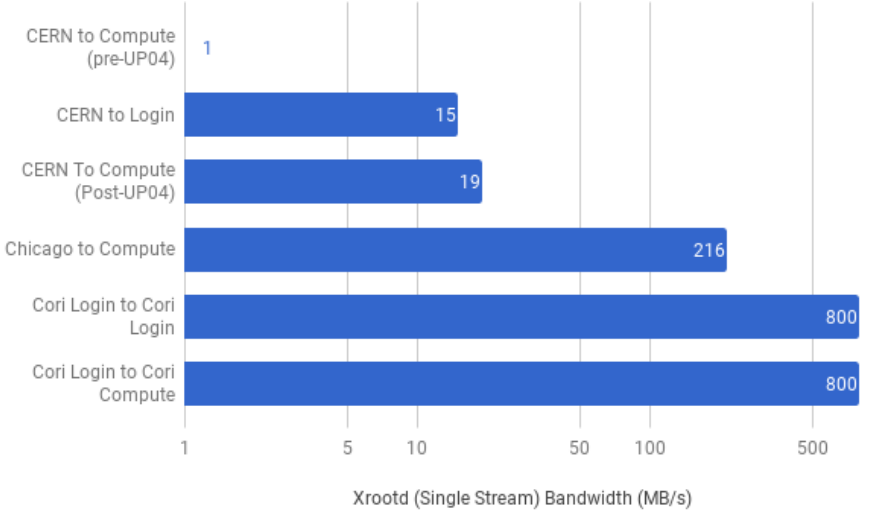
Atlas simulation **crash-test**, 1-CPU core, Jan Balewski
15 minutes simulation of 3 events, random requests of 60 releases of Atlas software, software and condition DB delivered via CVMFS, 32 concurrent tasks per node. 99 node job. **Failure rate: 1/60000**



WAN Networking to compute

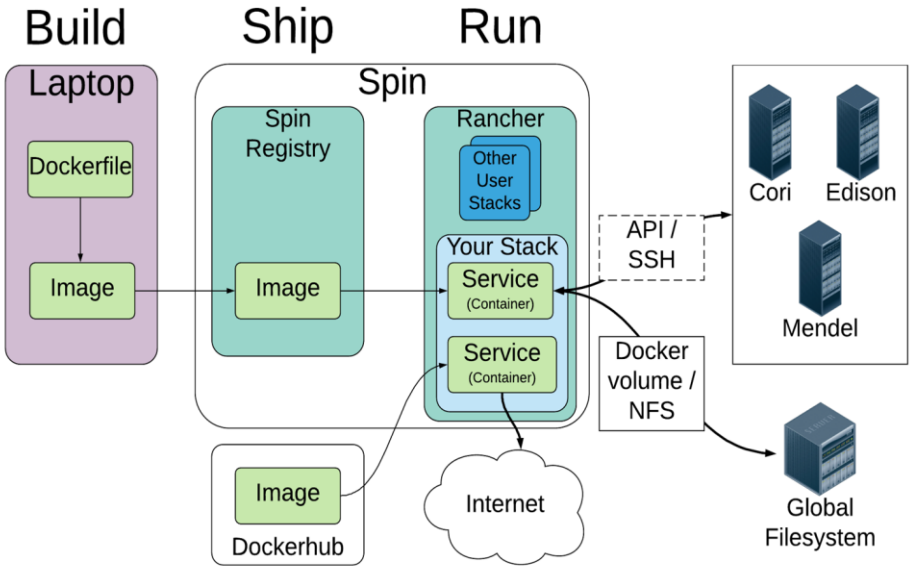
- **Cori compute nodes on 'Aries' high-speed network**
  - External traffic on Cray XC normally via 'RSIP' (limited performance)
- **'SDN' project: – first phase to replace RSIP with VyOS software**
  - iperf test 5 Gbs -> 26 Gbs
  - But TCP backlog drop on Aries affected transfers via some protocols (including xrootd) : fix in Aug 2017 OS upgrade
- **Xrdcp rates now exceed directly connected login nodes**



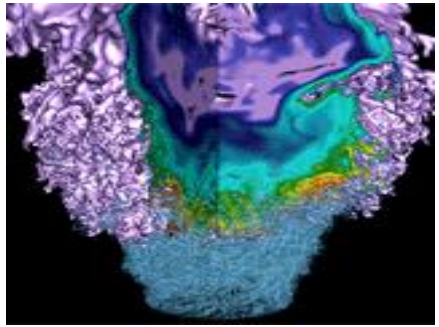


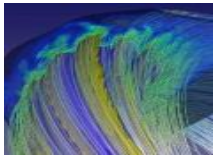
SPIN

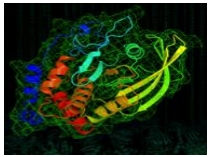
- **Container-based platform**
  - Can be used for scalable science gateways, workflow managers, databases and other edge services etc.
  - User-defined - minimal NERSC effort
- **Currently being commissioned:**
  - Early HEP projects with NERSC staff support
  - Squid servers; science gateways …

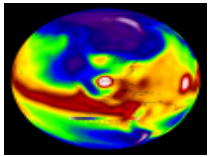


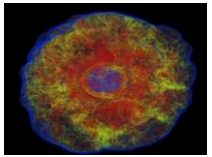
The coming future

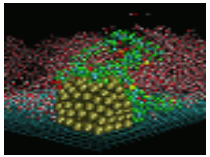












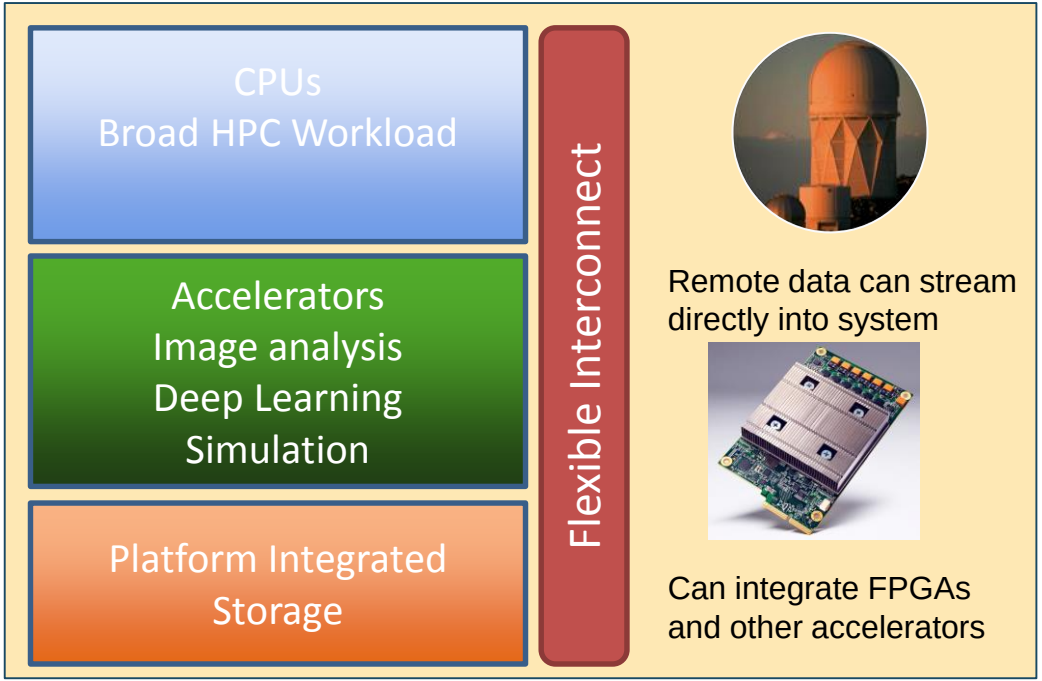
NERSC-9: A 2020 Pre-Exascale Machine

Capabilities

- 3-4x capability of Cori
- Optimized for both simulation and data analysis
- Looks ahead to exascale with specialization and heterogeneity

Features

- Energy Efficient architecture
  - Large amount of High-BW memory
  - High BW, low latency network
- Production deployment of accelerators for the DOE community
- Single-tier All Flash HPC filesystem



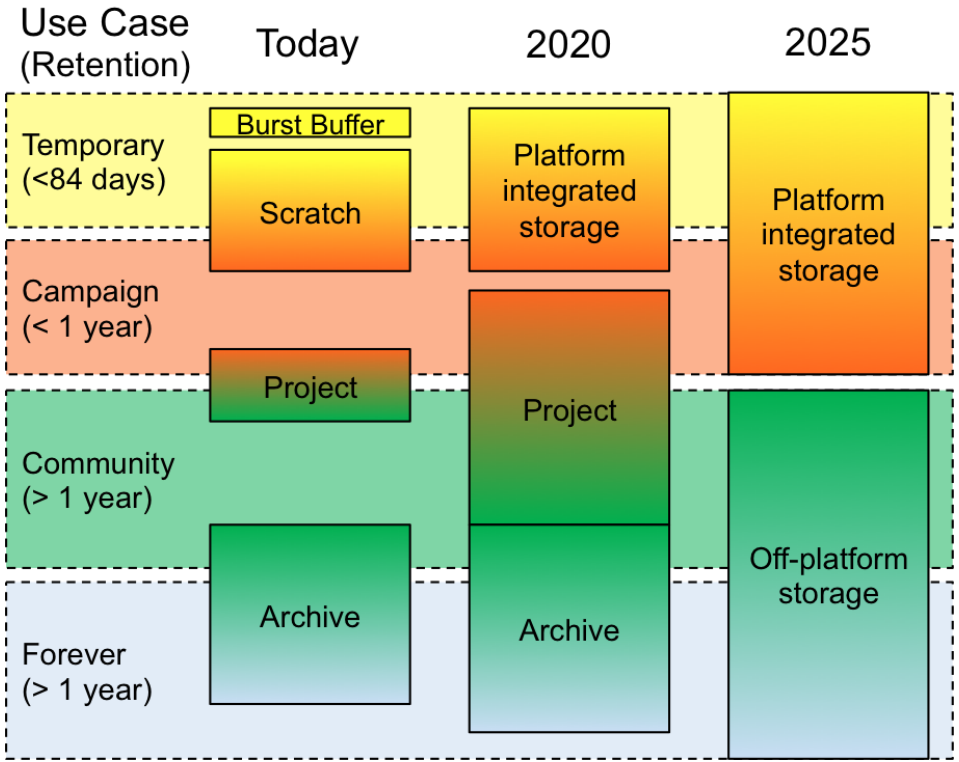
System will be announced in 2018

NERSC 'Storage2020' roadmap

All-flash parallel file system feasible for NERSC-9

> 100 PB disk-based file system

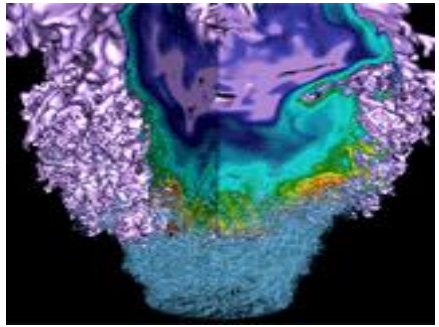
> 350 PB HPSS archive w/ IBM TS4500

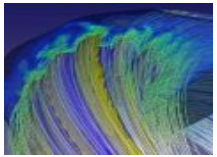


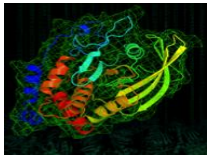
Conclusions

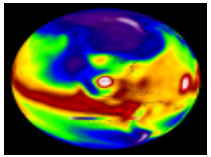
- **Many (current and planned) HEP experiments using HPC resources for *production* at NERSC (not just demos, stunts, or proof-of-concepts)**
  - CMS, ATLAS, LZ, Belle2, DayaBay, LSST-DESC, CMB-S4, DESI, NoVA …
  - Variety *of* ***different approaches taken and variety of Use Cases:*** MC Production; Reconstruction; Statistics. Enabling some workflows and scale/turn-around times that are not-possible with other resources
  - Also application porting (NESAP) and interactive machine-learning uses
- **Successes and also challenges. Experiments and NERSC have developed approaches and capabilities to run these workloads on 'big' HPC machines**
  - Machines that must still cater for broad workloads (>>90% non-hep-ex); be dense; power-efficient, manageable, and leading-edge
  - Technologies include Shifter; 'SDN'; DTNs; DVMFS; SPIN
- **Workflow and software barriers remain. Future brings increased support and new resources (N9, Storage2020, SPIN) but also architectural challenges**

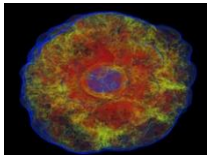
Backups

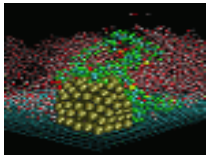








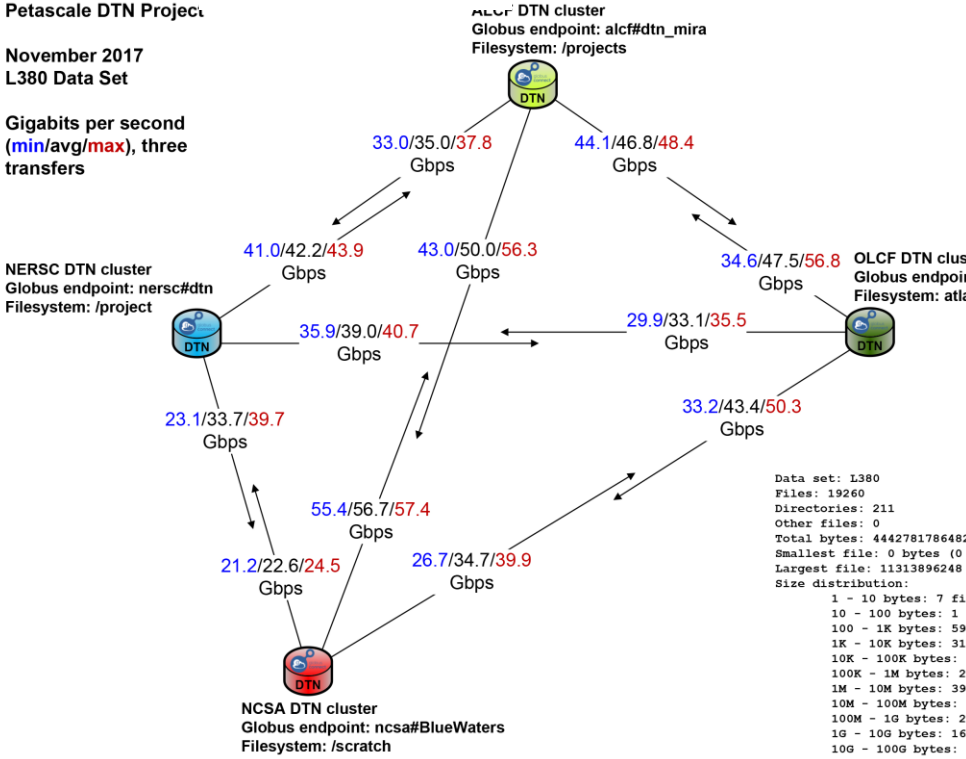




Production transfers to DTN

Eli Dart *et. al.*

- **‘Petascale’ project with ESnet and others driven performance from ~6-10 Gbs to ~20-~40Gbs)**
  - For ‘real’ datasets
  - Onto NERSC project filesystems (via DTN nodes)
- **HEP experiments (e.g. ATLAS) using FTS/rucio to pure-GridFTP DTN endpoints**
- **Expand testing to FNAL, BNL**

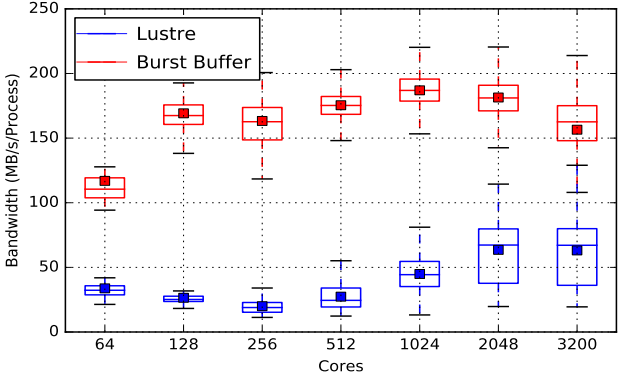


Burst Buffer

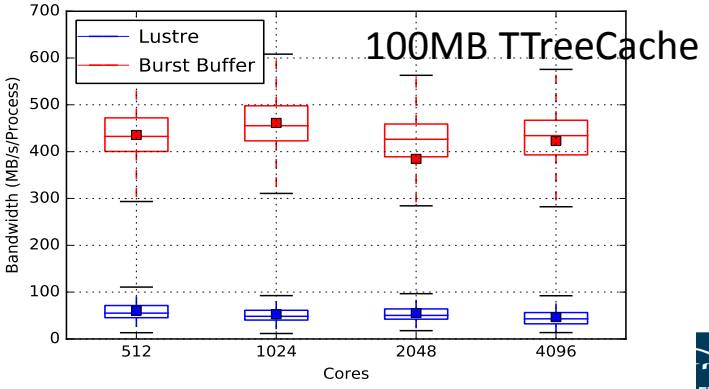
WB, Steve Farrell, Vakho Tsulaia et. al

- **Burst Buffer: 1.8 PB all-flash**
  - DataWarp: on-demand posix filesystems
  - Benchmark peak:, ~1.7 TB/s (read/write)
- **Potential performance gains for many I/O heavy workloads in HEP experiments shown at CHEP 2016**
- **Outperforms Lustre and scales well (and both are good for bulk I/O)**
- **Many production and large-scale workloads using BB**
  - But not so much by HEP-ex production … insufficient gain for those workloads – possibly due to other overheads

Derivation (xAOD->xAOD) in AthenaMP



QuickAna on 50TB xAOD dataset

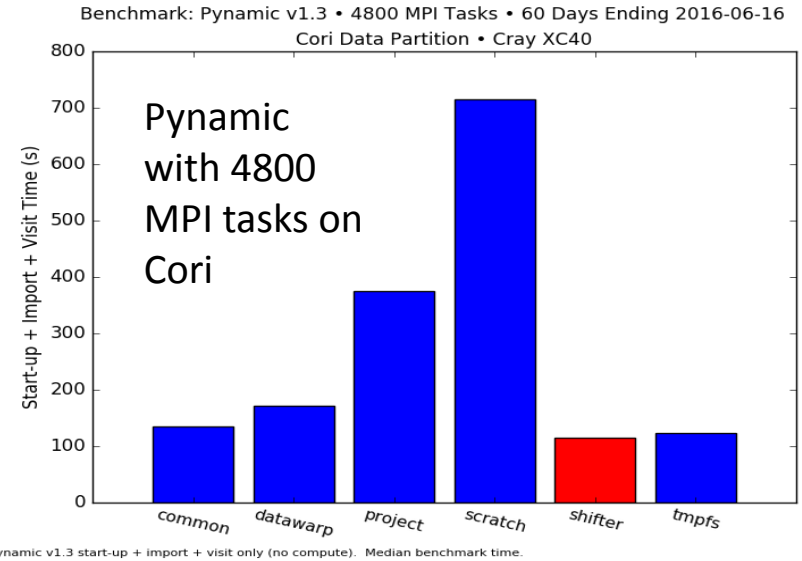


Shifter

Doug Jacobsen,
Shane Cannon *et. al.*

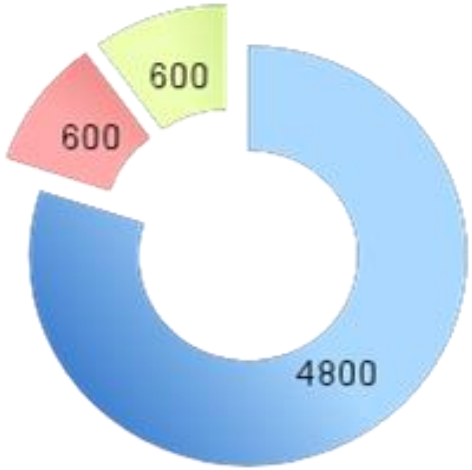
- **Crays run CLE OS (modified SLES)**
  - Linux codes should compile fairly easily but packages can be different – containers are a solution
- **Shifter allows users OS stack**
  - Imports docker (or other) images
- **Integration with HPC software and architectures**
  - MPI and other system libraries, integration with workload managers,
  - Volume mount NERSC filesystems
  - NERSC retain control of privilege
- **In use by ATLAS, CMS and numerous small HEP experiments**
  - Recipes including MPI
  - Example containers for HEP
- **Also has benefits for shared library loading**

Plot: Rollin Thomas



How time on NERSC big machines is allocated

Allocated Hours 2017 (Millions)



DOE Mission Science	80%
ALCC (computing challenge)	10%
Directors Discretionary	10%

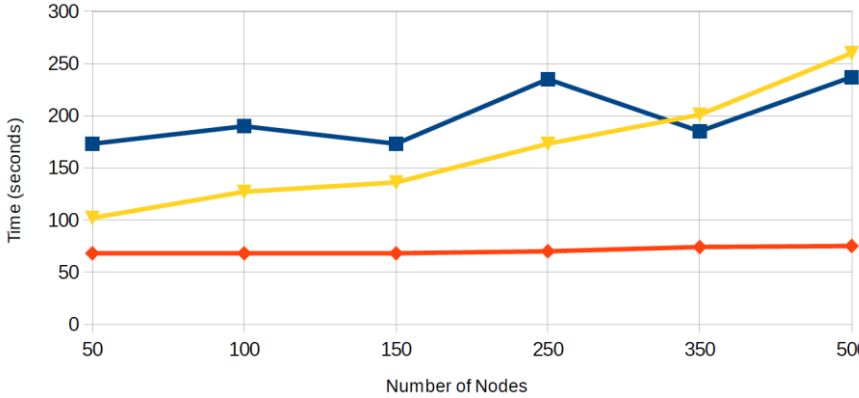
- Time (mostly) allocated by DoE science program managers
  - ~15% HEP (including lattice, cosmo etc.)
  - Recently large allocations to LHC
- Yearly allocations though some hope/plan of being able to allocate longer ones
- Scratch and project disk storage 'included' at ~10 TB level though larger on request
  - As is archive /HPSS
  - Some buy sponsored storage (e.g. Daya Bay)
- 'PDSF' cluster is different - 'owned' by those HEP experiments with fairshare division
- Machines popular – little opportunistic idle time. But backfill possible (esp. for small, short jobs) due to e.g. draining for large jobs

Software - CVMFS

Lisa Gerhardt, Vakho Tsulaia …

- **Historically NERSC systems have not been keen on fuse**
- **One approach is to ‘stuff’ cvmfs into a container:**
  - unpack cvmfs; removing duplicates (with e.g. ‘uncvmfs’) and build SquashFS image
  - Working in production for ATLAS and CMS
  - Now users can build even these big images – NERSC loads to shifter

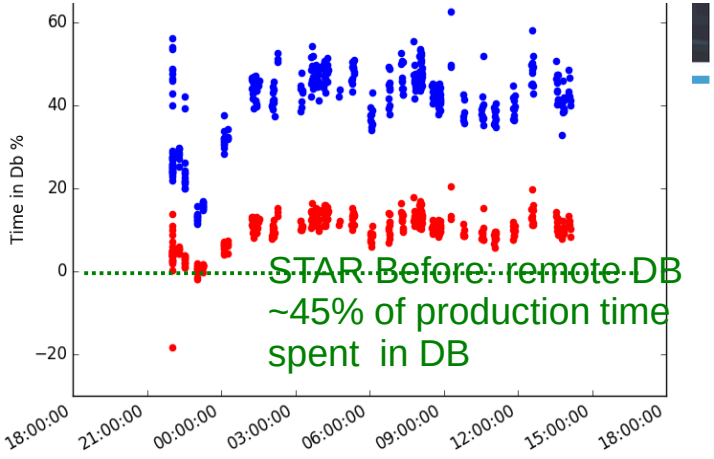
AthenaMP startup time
Shifter; Burst Buffer; Lustre

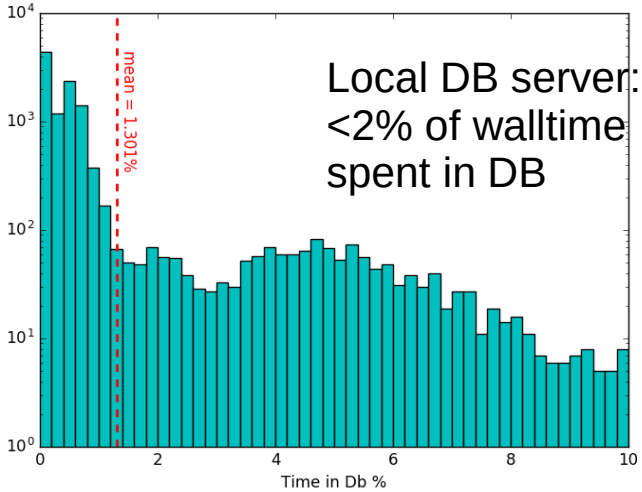


Small file I/O

- **Burst buffer and Lustre recent enhancements for small file I/O**
  - DVS client side caching in BB
  - Multiple meta data servers (DNE) for Lustre
- **Also shifter perNodeCache**
  - Temporary xfs filesystem all metadata on WN
  - E.g. used for STAR read/only copy of mysql database on compute nodes
  - CMS use for madgraph jobs

Shane Cannon Mustafa Mustafa et.al.





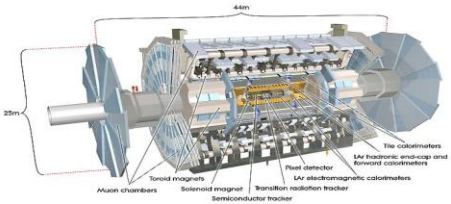
Rollin Thomas et. al.

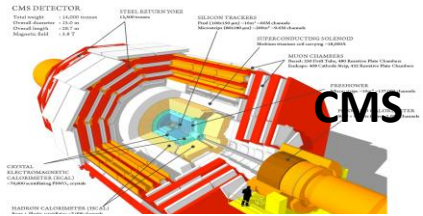
NESAP: Software

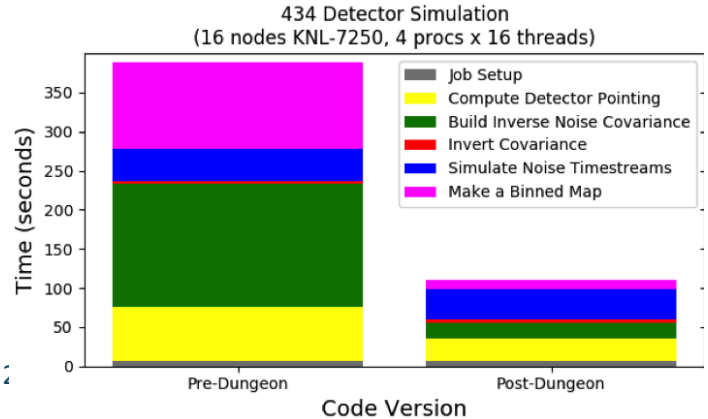
NERSC Exascale Science Applications Program

- **Extended NESAP program to projects processing experimental science data: "NESAP for Data"**
- **Had call: 4/6 teams chosen were HEP**
  - Teams get postdoc at NERSC
  - And vendor collaboration (dungeon sessions), extra support from NERSC.
- **Plan to continue NESAP for Nersc-9 with "data" apps from the outset**

ATLAS







Recent TOAST Dungeon Improvement Ted Kisner,

Workflows: and SPIN

- **Now deploying container-based platform (SPIN) to create scalable science gateways, workflow managers, and other edge services with minimal NERSC effort**
- Ultimately seek to provide software/API for (e.g.) data transfer/sharing, migration between file system layers, scheduling, usage query, job/workflow status (Superfacility API)
  - Build on existing best practice