



Lawrence Berkeley
National Laboratory



HPCOR 2014

System Configuration (D1SA)

Co-Chairs:
Clay England (OLCF) &
Jason Hick (NERSC)

Contributors

- *Andrew Cherry (ALCF)*
- *Cory Lueninghoener (LANL)*
- *Curt Canada (LANL)*
- *Robin Goldstone (LLNL)*
- *Jim Silva (LLNL)*
- *Jason Hick (NERSC)*
- *Nick Wright (NERSC)*
- *Clay England (OLCF)*
- *Chris Beggio (SNL)*
- *Bob Ballance (SNL)*
- *Glenn Lockwood (SDSC)*
- *Eli Dart (ESnet)*

What is a data analytics system and what are the hardware characteristics?

- A new area for HPC which needs a flexible system
- Analytics is an HPC problem due to scale
- Similar data movement issues to exascale and data intensive workloads both are in need of software to move through the storage hierarchy
- Large amount of memory/core
- Fast I/O and Persistent storage
- Require a burst buffer (NVRAM close to compute)
- Must support wide variety of I/O (IOPs to bandwidth)

Are HPC and data systems different or the same?

- There is no one size fits all
- There are similarities software/hardware/storage to HPC and analytics
- The more specialized or heterogeneous the HPC system is the more difficult it is to use for analytics.
- Does App Readiness have someone focused on visualization and analytics?

What percentage of resources should be allocated to compute vs. I/O &/or storage?

- Begin with memory/core
- Then SSD/persistence
- Next PFS/network & I/O
- From there figure out how much CPU you can afford.
- Could be different at each center with workload analysis.
- Need for better tools to monitor data/storage and I/O

What opportunities exist for productive collaborations among DOE HPC centers?

- Workload analysis for data analytics, identify and share this analysis
- Develop common workload analysis tools
- Demand for multi-facility workflows is increasing especially between experimental and HPC facilities
- Collaboration on using flash with parallel file systems
- NERSC, SNL, and LANL developing burst buffer together
- Storage and network QoS locally and between facilities
- Establishing performance expectations for data ingest and export

What are the biggest challenges and gaps between what you can do today and what will be required in 5 - 10 years?

- Software stacks the facilities are looking at adopting are not written by or for the HPC community
- Data analytics is new to HPC community/facilities. How do we manage/schedule this resource? What are the success metrics?
- Custodial/Policy issues with increased amount of data
- Differences between security zone and account between facilities.
- Data movement cost will continue to overshadow computational capabilities

Top Findings

Opportunities

- Definition and comparison of workload analysis
- Flash with PFS
- Best Practices
- Burst Buffer
- Storage and Network QoS

Best Practices

- New to HPC
- Workload defined architecture

Challenges

- Analytics software stacks aren't HPC today.
- Mutli-site analysis is today untenable.
- Data movement
 - Inter-node
 - Between facilities