I/O Performance of Virtualized Cloud Environments

Devarshi Ghoshal,

Richard S. Canon, Lavanya Ramakrishnan

U.S. DEPARTMENT OF ENERGY | Office of Science

NERSC National Energy Research Scientific Computing Center

BERKELEY LAB Lawrence Berkeley National Laboratory

Data Volumes are Increasing

- Data is a critical component of scientific process
  - Joint Genome Institute projects about 2PB/year
  - Large Hydron Collider (LHC) projects 16PB/year
  - Large Synoptic Survey Telescope (LSST) projects 6PB/year of raw data
- Similar problem: Internet data
  - Terabytes to petabytes of data/day
  - Cloud is resource platform for these applications
    - MapReduce/Hadoop
    - On-demand virtual machines for processing

What is a Cloud?

- Cloud Infrastructure as a Service (IaaS)
  - Provision bare metal or virtual machine resources
  - Pay-as-you-go model
    - different quality of service levels might be available
  - User responsible for complete software stack
  - Storage options: Elastic Block Store, Ephemeral local disk, S3

Magellan

Determine the appropriate role for commercial and/or private cloud computing for DOE/SC midrange workloads

Approach

- Deployed a distributed test bed at Argonne and NERSC to explore the use of clouds for mid-range scientific computing.
- Evaluated the effectiveness of cloud models for a wide spectrum of DOE/SC applications.

Goal

- Benchmarking I/O performance over different cloud and HPC platforms to identify major bottlenecks
  - Network and I/O related
    - Identified through prior Magellan research and other related efforts
- I/O Intensive
  - Low Disk Latency
  - High Bandwidth
  - High Speed Interconnection Network
  - Scalability
- Specialized infrastructure at supercomputing centers compared to cloud's commodity infrastructure

Related work

- Disk I/O and Throughput on Amazon EC2
  - I/O benchmarking for different storage options available on Amazon's EC2 cloud infrastructure
  - *Not sufficient data to understand impact on HPC apps*
- Performance Analysis of HPC applications on Amazon EC2
  - Focus on determining the performance of HPC applications on cloud
  - *No work on the effect of virtualization on I/O*

NERSC Configuration

- NERSC Machine Description
  - 720 node IBM iDataPlex cluster
    - 40 nodes for I/O benchmarking
  - Each node has two quad-core Intel Nehalem processors running at 2.67 GHz, 24 GB RAM
  - 4X Quad Data Rate (QDR) Infiniband Technology
  - Batch Queue system
    - GPFS, peak performance of 15 GB/s
  - VM instance type- c1.xlarge

Amazon EC2 Configuration

Instance type	API name	CPU Family	EC2 Compute Units	Memory (GB)	Local Storage (GB)	Expected I/O Performance
Small	m1.small	Intel Xeon E5430	1	1.7	160	Moderate
Large	c1.xlarge	Intel Xeon E5410	20	7	1690	High
Cluster-compute	cc1.4xlarge	2 x Intel Xeon X5570	33	23	1690	Very High

Benchmarks

- IOR (Interleaved or Random) benchmark
  - For testing performance of parallel file systems
  - Read and Write Performance
  - I/O type: Direct, Buffered
  - Low benchmarking overhead
- Timed I/O Benchmark
  - Performance measurements for a particular duration
  - Measurement results at specific time-intervals

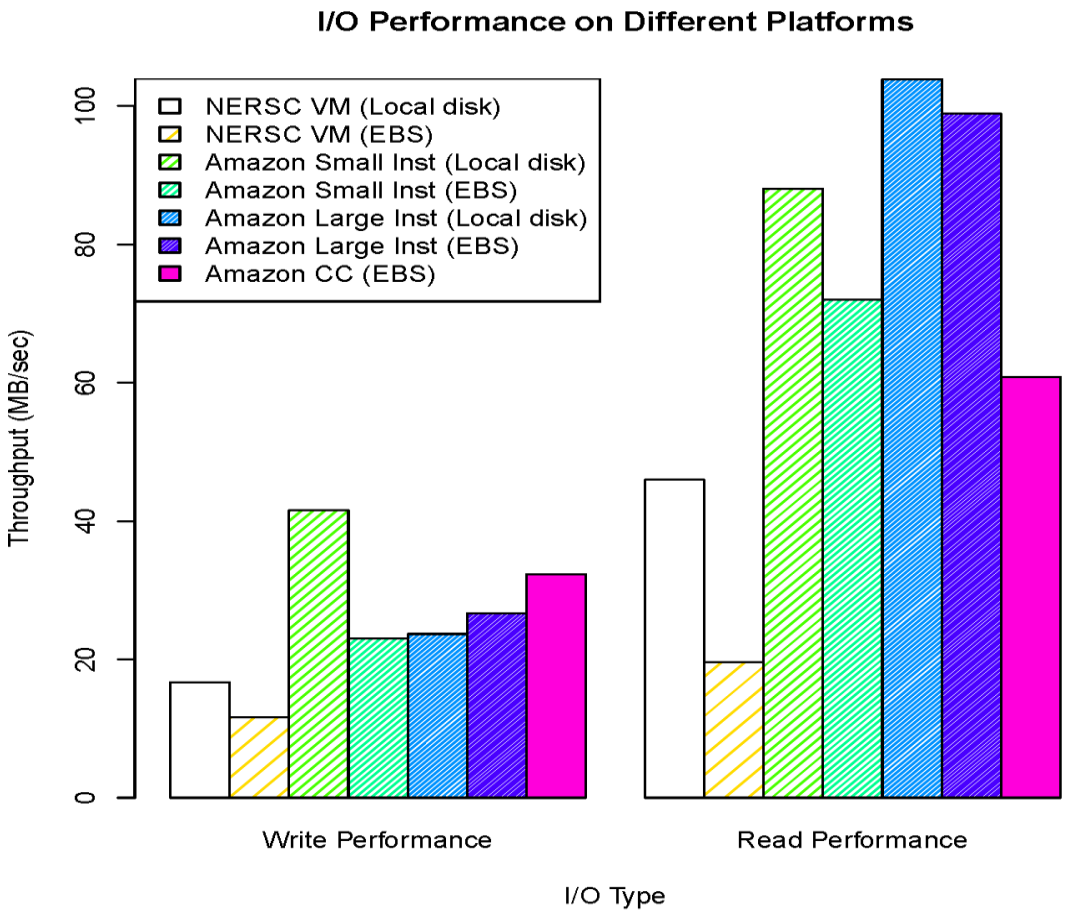
Benchmark Parameters

- Types of I/O
  - Direct I/O vs Buffered I/O
- Virtual Machine Instance Types
  - Small, large and Cluster-compute instances
- Storage Devices
  - Local stores (ephemeral), EBS volumes
- Location of Instances
  - US-East and US-West regions
- Shared File-system
  - GPFS vs EBS volumes
- Time of run

Evaluation Summary

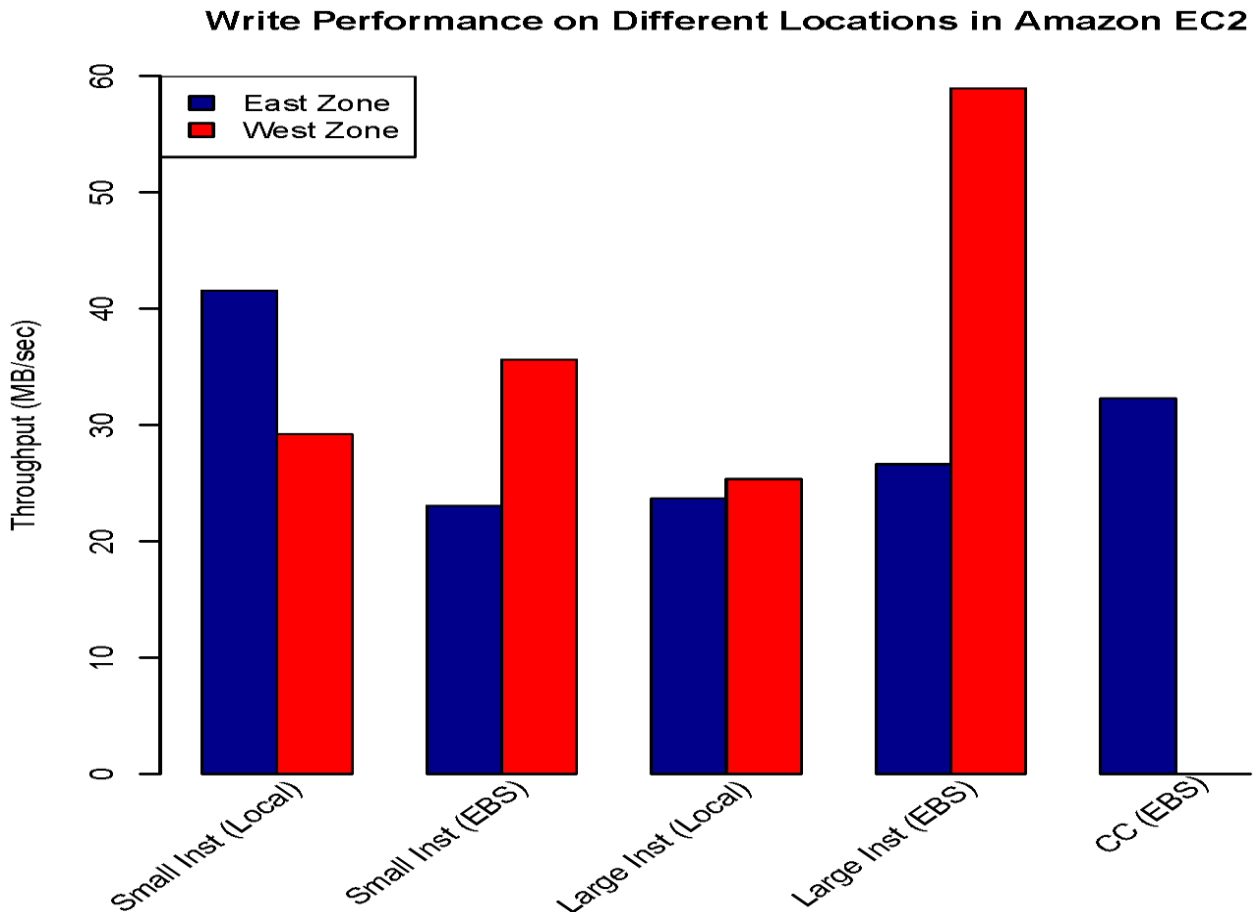
- Buffer Caching
- **Comparison across all Platforms**
- Effect of Regions on Amazon
- **Multi-node Shared File-system**
- **24-hour Tests**
- **Large Scale Tests**

I/O Performance- Summary



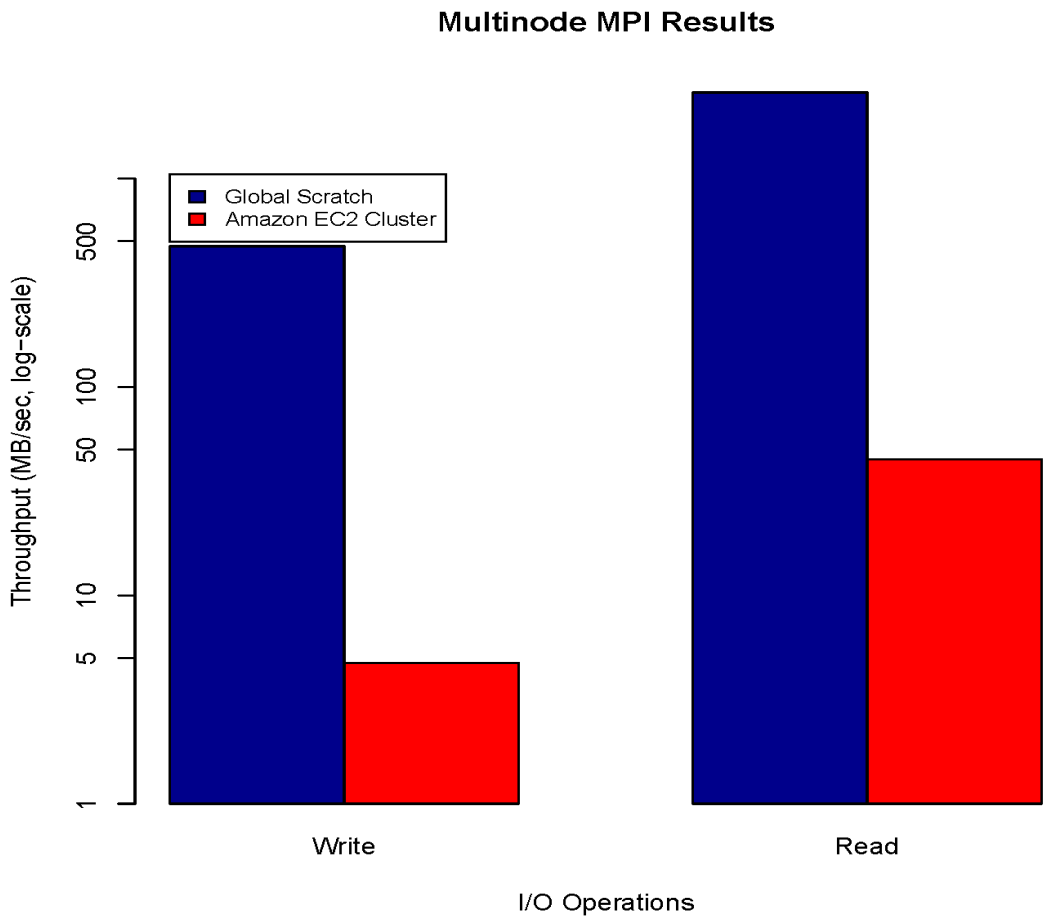
- Network bandwidth plays an important role in determining I/O performance.
- EBS performance is better on CC instances, due to 10 Gigabit Ethernet network.

Amazon EC2 (Write Performance)



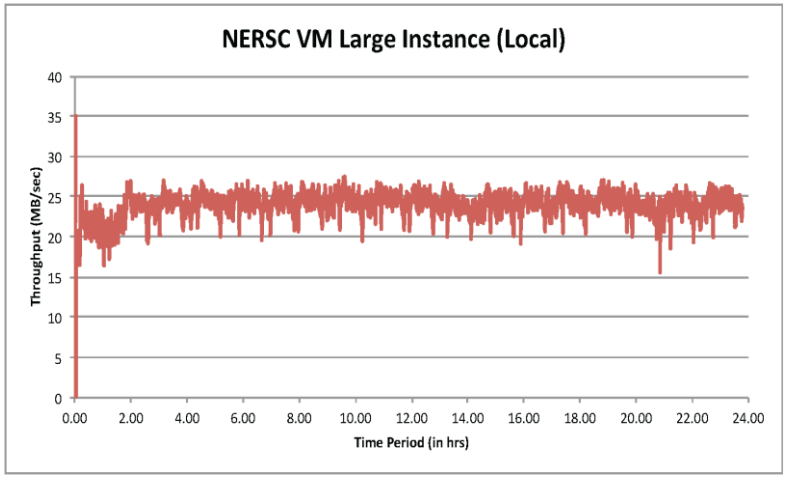
West-zone instances outperform the East-zone instances in most cases

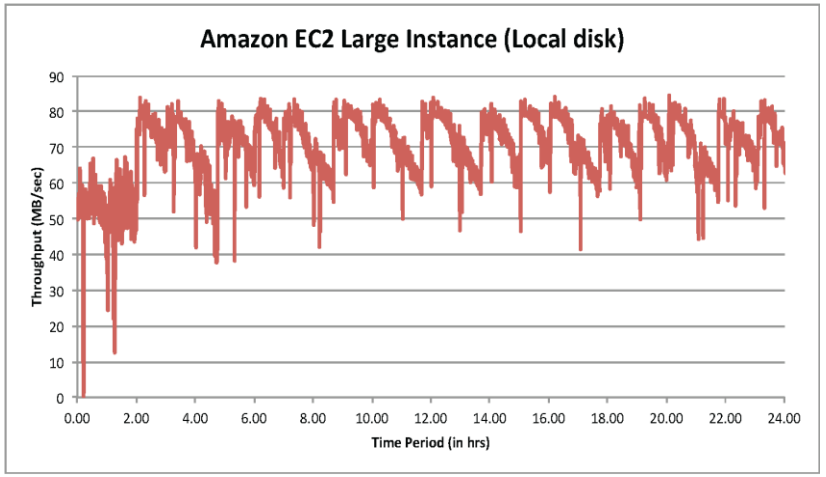
Shared Filesystem

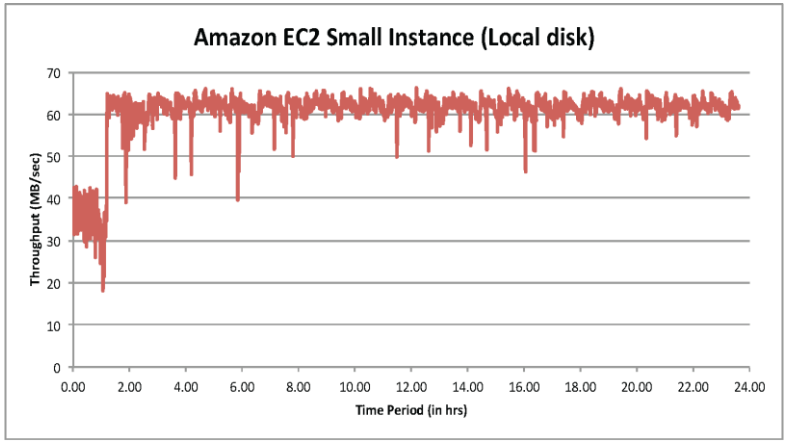


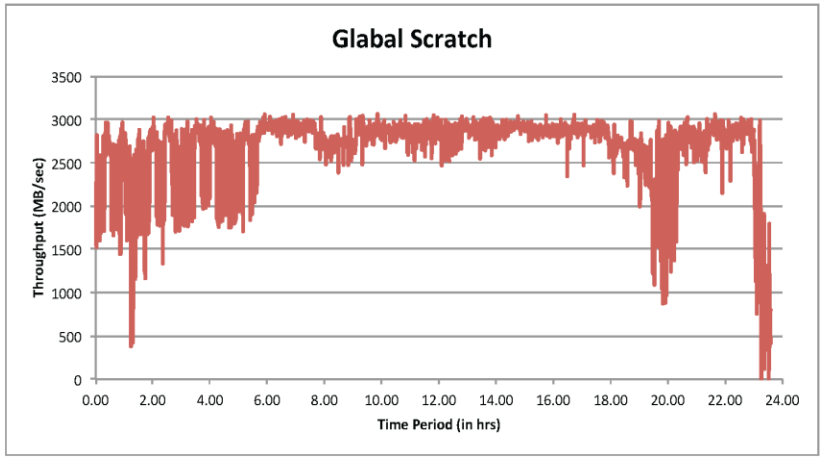
High resource contention degrades the EBS performance severely.

Timed Benchmark (Write)



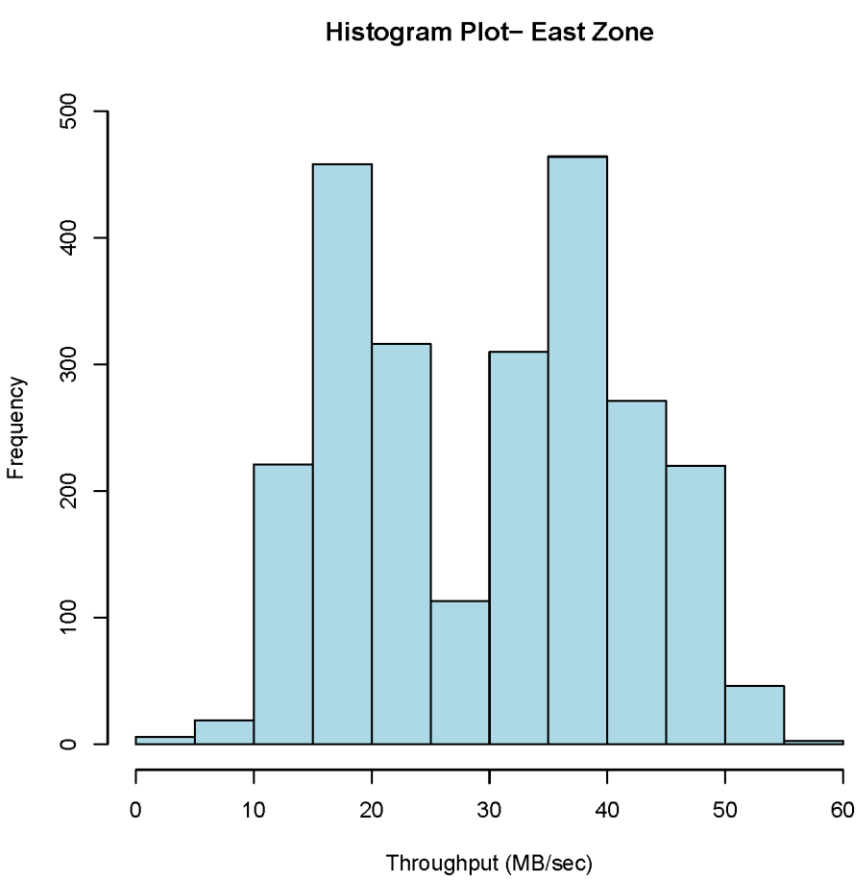


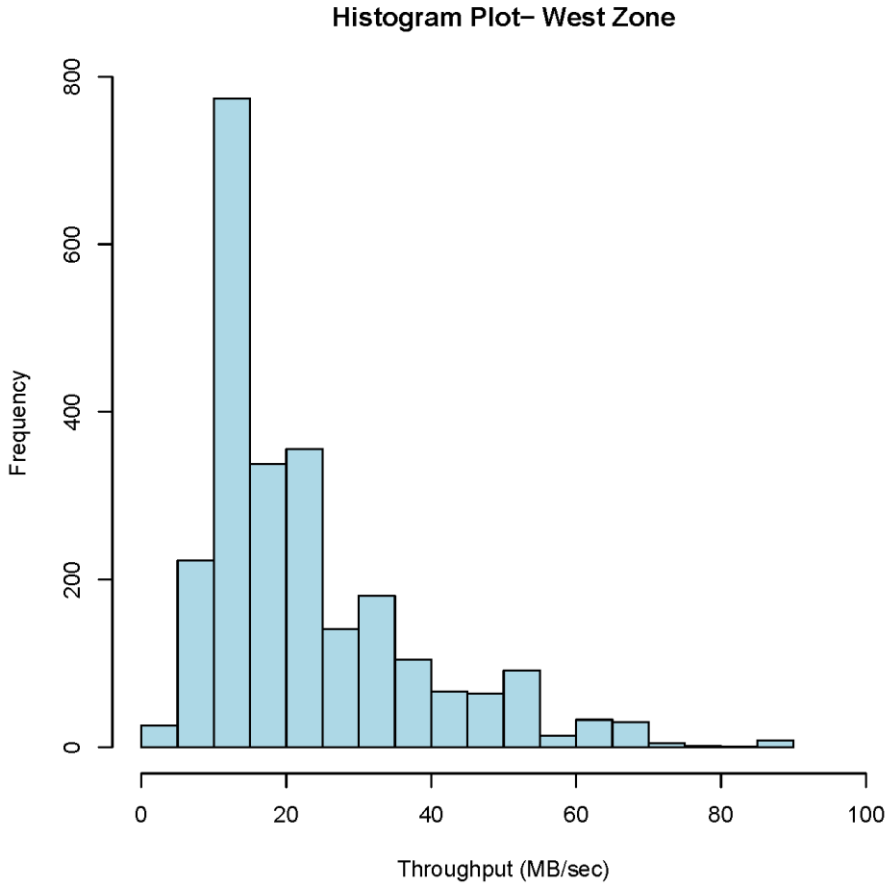




Occasional performance drop may be attributed to the underlying shared resources.

Performance Distribution





West-zone performance varies a lot compared to East-zone, but the average is lower.

Impact on applications

- Performance impact on I/O
- Suitable trade-offs between EBS and local disk
  - persistence vs performance
  - cost is another factor
- Application design needs to consider
  - performance variation
  - lack of high performance shared file system

Conclusions & Future Work

- Performance in VMs is lower than on physical machines
  - Clouds are not yet ready for data-intensive applications with high-performance requirements
- I/O performance on local disks is better than on EBS volumes
  - Local disks are ephemeral devices
  - Local disks are also not suitable to MPI (most commonly used in HPC applications)
- Timed results show substantial variations
  - Further investigation required

Acknowledgements

- US Department of Energy DE-AC02-05CH11232
- Krishna Muriki, Iwona Sakrejda, David Skinner

Questions?

Other related events at Supercomputing –

- At Lawrence Berkeley Booth
  - Science in the Cloud? Busting Common Myths about Clouds and Science Tue Nov 15 at 10:30 am
  - What do Clouds mean for Science? Experiences from the Magellan Project Tue Nov 15 at 11:15 am
- Papers
  - Evaluating Interconnect and Virtualization Performance for High Performance Computing, PMBS Workshop, Sun Nov 13 at 9 am.
  - Riding the Elephant: Managing Ensembles with Hadoop, MTAGS Workshop, Mon Nov 14 at 4:30 pm