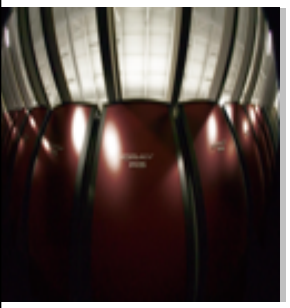


Sandia National Laboratories

Exceptional service in the national interest

Output Performance Study on a Production Petascale filesystem

Bing Xie, Jeff Chase, David Dillow, Scott Klasky, **Jay Lofstead**, Sarp Oral, Norbert Podhorszki

Duke University, TheDillows.Org, Oak Ridge National Lab

Scalable System Software
Sandia National Laboratories
Albuquerque, NM, USA
gflofst@sandia.gov

June 22, 2017

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Performance, Interference, Locality

Sandia National Laboratories

- Experiment with various configurations
- Determine what mitigations might work
- Evaluate various sizes
- Uses IOR to drive tests
- Generally 1 proc per node to 1 OST with combinations

Terms and Standards Used

Sandia National Laboratories

- Bandwidth is measured as MB/s per client node.
- Aggregate Bandwidth, measured in MB/s, is bandwidth summed across all client nodes in an instance.
- Effective Aggregate Bandwidth (EAB) is aggregate bandwidth normalized to the peak bandwidth achievable from the number of targets written in an instance under a given set of parameters.

Titan File System Characteristics

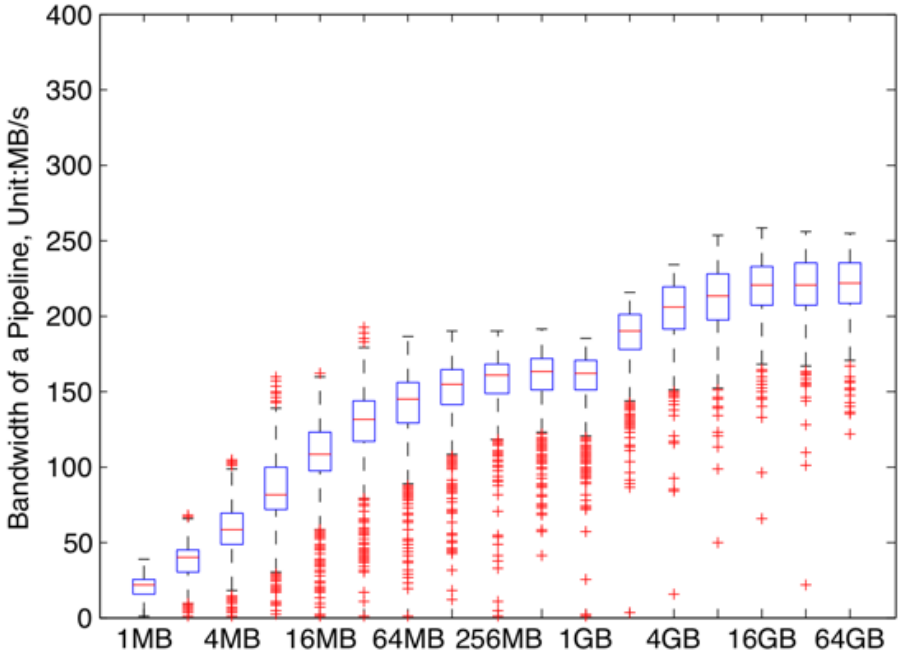
Sandia National Laboratories

- Two configurations tested

File Systems	Service Time	Partitions	Routing Policy	I/O Nodes	Osiers	OSTs
Spider	Jan.2008–Dec.2013	4	fine-grained	192	192	336 × 4
Spider 2	Nov. 2013–present	2	fine-grained	432	288	1008 × 2

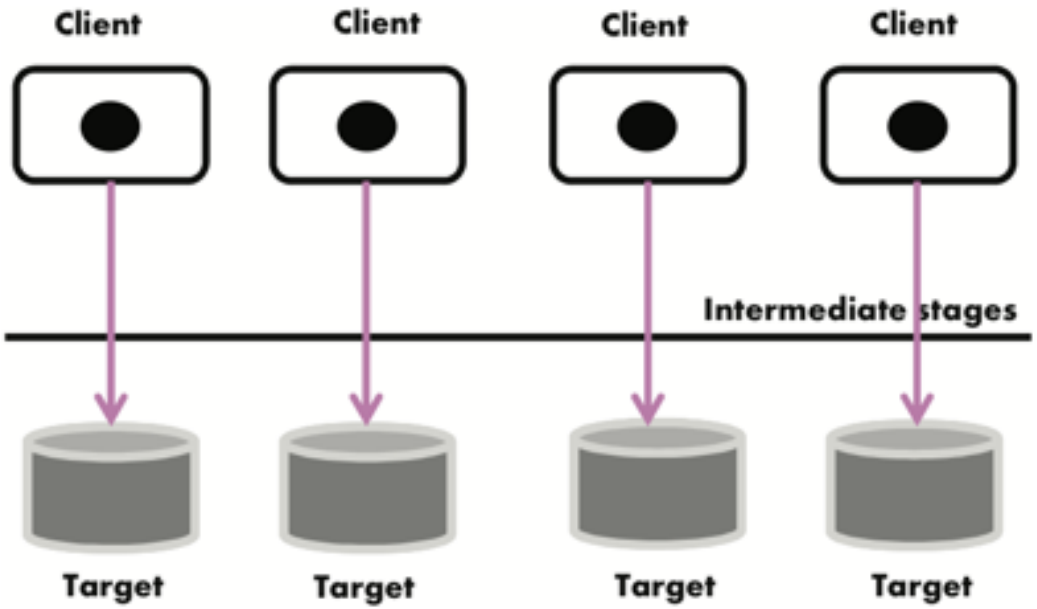
Single Process/Node to Single OST

- March to July 2013 on Spider/Widow1 to a single file
- Additional client processes did not help obtain full bandwidth (max 5% improvement)
- 3-5% are significantly slower (intermittent interconnect interference).



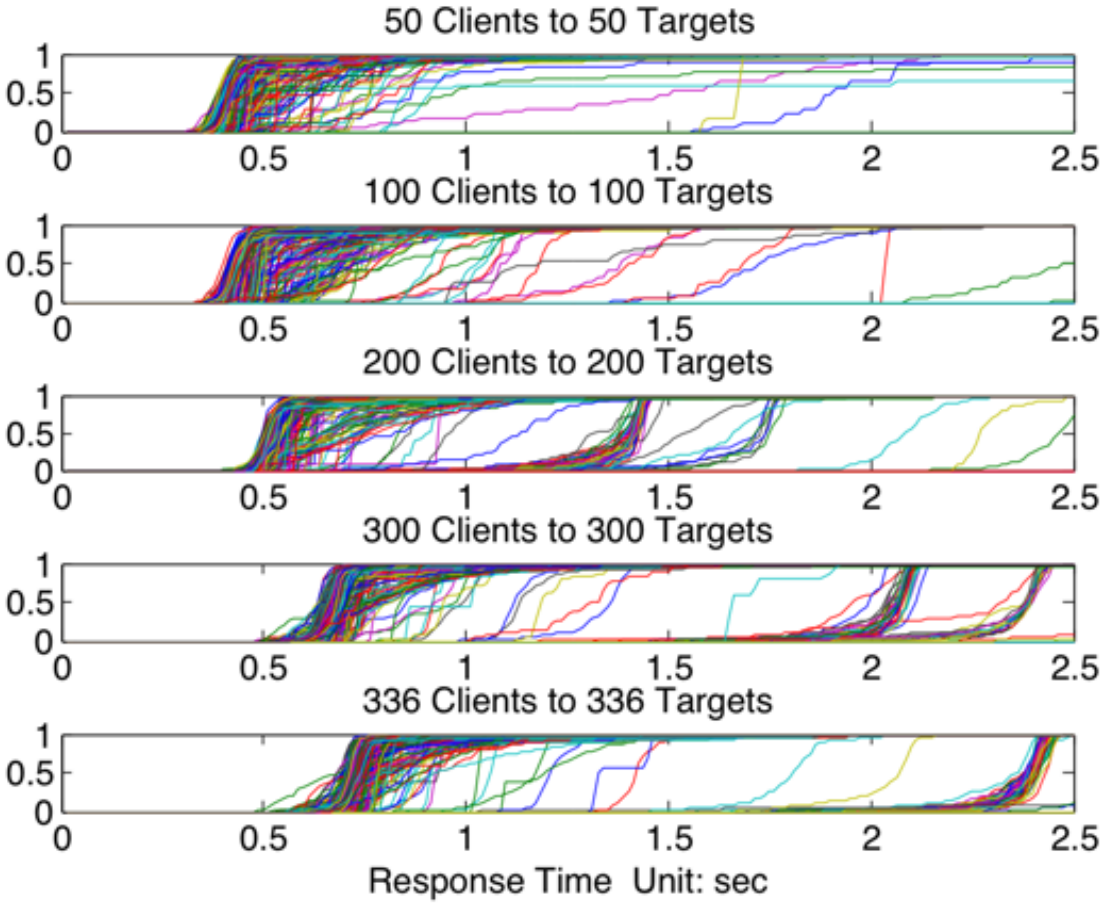
Many Pairs Setup

- February to July 2013 on Spider/Widow 1
- Each node runs 1 client writing 64 MB to an unstriped file on a single, different OST
- Output synchronized
- Measure aggregate bandwidth and completion time for each pair.



Many Pairs Results

- Lots of stragglers, particularly as the number of clients (and OSTs) increase.



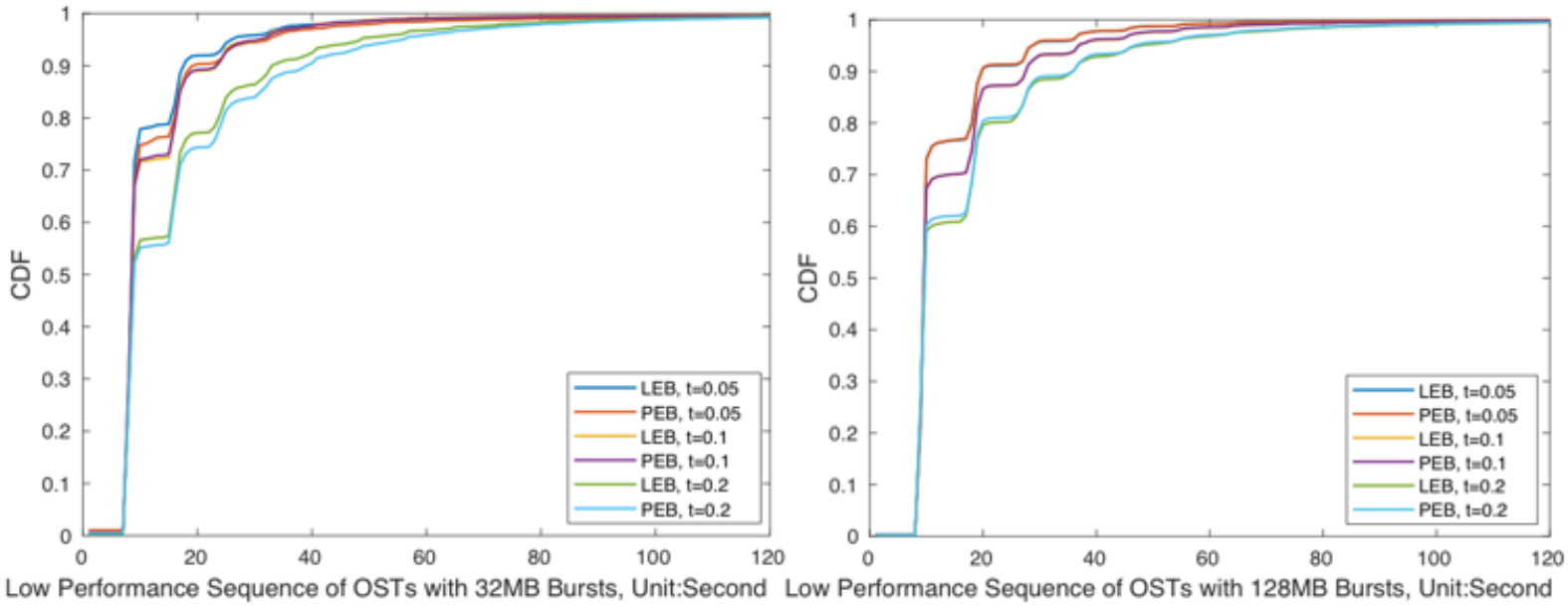
Testing Persistence of Stragglers

Sandia National Laboratories

- January to February 2017 on Titan/Atlas2 with 1008 OSTs.
- 126 clients against 126 OSTs round robin 32 MB and 128 MB bursts.
- Duration is 0.5-1.7 hours for run with 7-15 seconds between runs
- Normalized values
- Lag Effective Bandwidth (lower bound)
- Pair Effective Bandwidth (upper bound)
- 99.5%+ of scores in 0.4-0.9 range

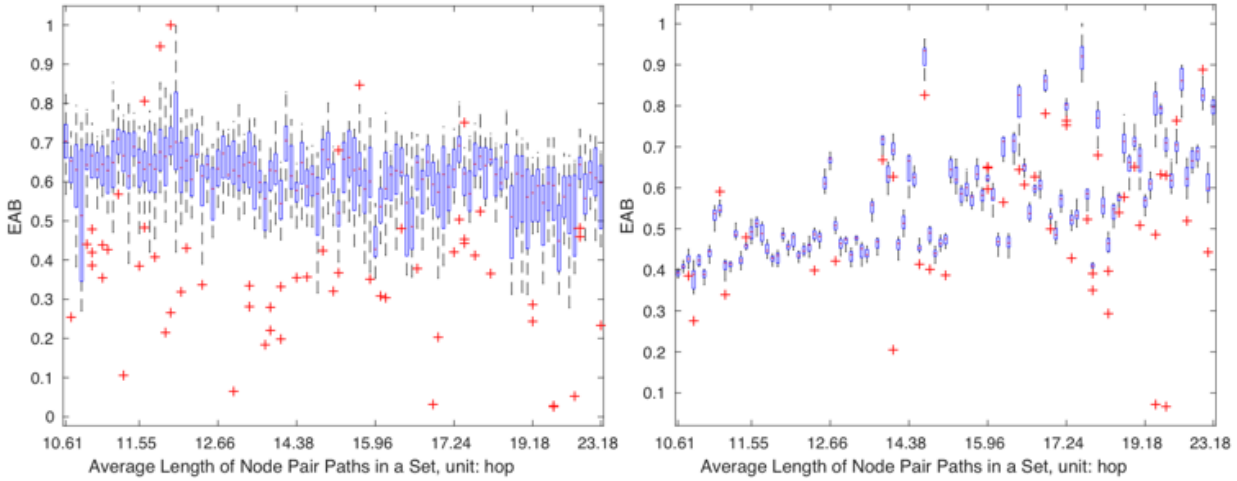
Individual Component Variability

- 96% of 32 MB and 100% of 128 MB “slow” targets return to normal within 2 minutes.
- Any node showing high performance will slow within 10 minutes.



Node Locality

- May to June 2015, 95 sets of 16 MB bursts (left) and 256 MB bursts (right)
- X-axis is sorted average node distance
- Y-axis is distribution of measured aggregate bandwidths
- Small burst dominated by transient conditions
- Large bursts impaired by denser node sets



Conclusions

Sandia National Laboratories

- We find that a small proportion of storage targets (< 20%) are straggling at any given interval, but that stragglers are transient: over time, any target may appear as a straggler for some intervals. Stragglers throttle the write pipelines, limiting striping bandwidth and reducing the benefits of paral- lelism.

- As configured on Titan, the Lustre write pipelines do not allow a single client to obtain the full bandwidth of a storage target. The results suggest that in the ideal case each client writes to multiple files spread across multiple targets, with multiple clients per target.

Conclusions

Sandia National Laboratories

- The I/O performance delivered on Titan is highly variable. Our study suggests that historical performance data and monitoring do not enable adaptive middleware to locate “good spots” in the supercomputer or in the file system. Local performance behavior is transient and unpredictable.

- Delivered aggregate output bandwidth is sensitive to location (density) of a job’s compute nodes for large bursts, under a static node-to-router mapping policy adopted by Titan in its internal network configuration.

Questions?

Sandia National Laboratories

Jay Lofstead

gflofst@sandia.gov