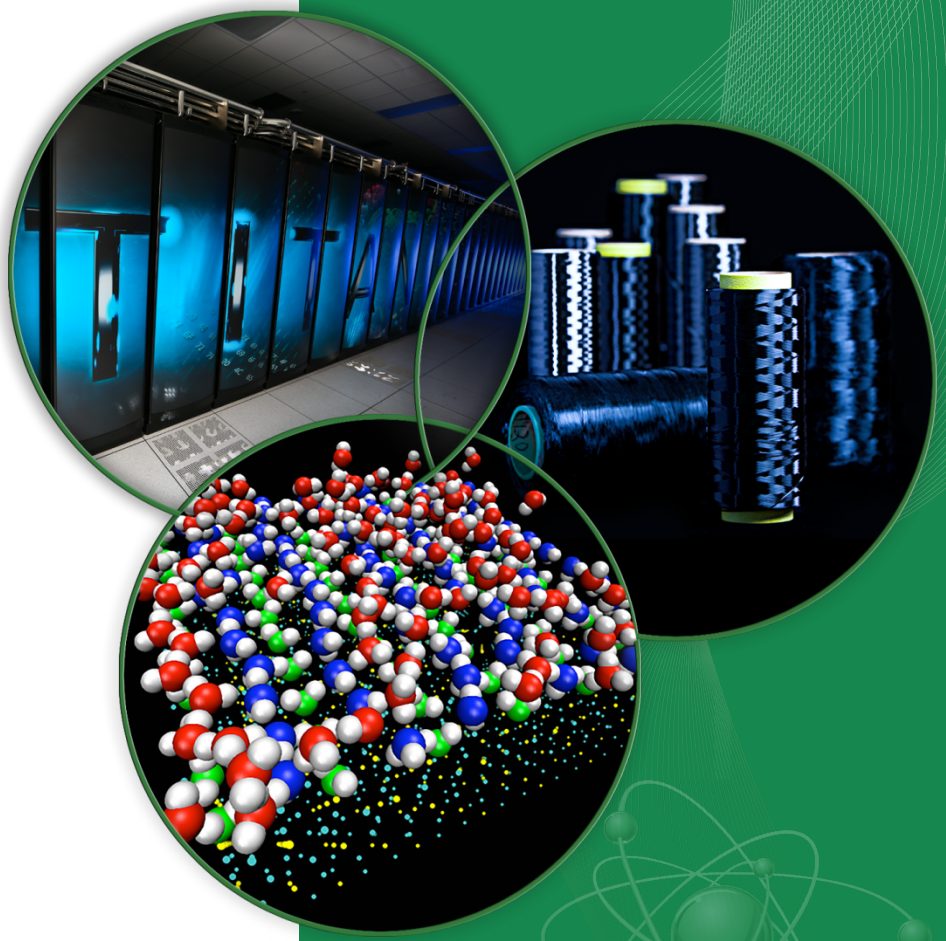
A More Realistic Way of Stressing the End-to-end I/O System

Verónica G. Vergara Larrea
Sarp Oral
Dustin Leverman
Hai Ah Nam
Feiyi Wang
James Simmons



CUG 2015
April 29, 2015
Chicago, IL

ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

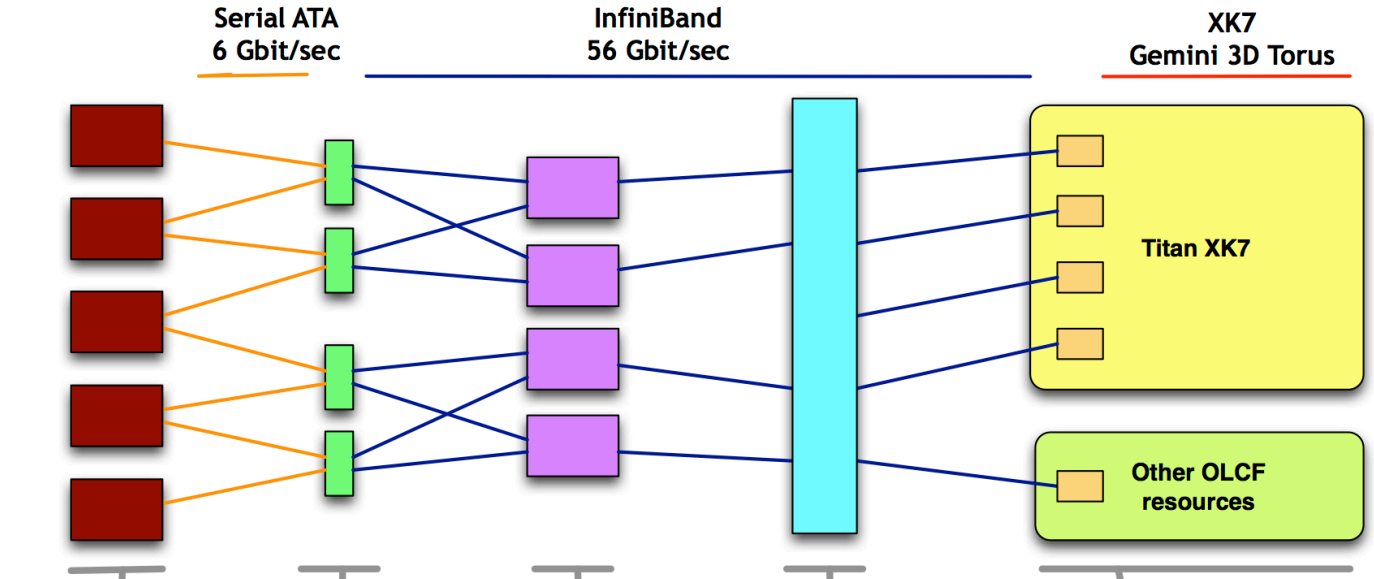
Outline

- Background
- The OLCF end-to-end I/O path
- Experiences upgrading Lustre
- I/O test harness
- Testing Methodology
- Testing on Titan
- Summary & Future Work

Background

- OLCF is home to Titan
  - Cray XK7 system with 18,688 nodes
  - #2 in the TOP500 list (November 2014)
- Spider 2: center-wide Lustre file-system
  - 32 PB, 1TB/s
- Testing at the extreme scale
  - Rigorous acceptance testing during design, acquisition, and deployment
  - Upgrading a production file system also requires similar levels of testing
  - Vendors lack the capabilities to test at large-scale
- Popular testing practices of file systems include synthetic benchmarks
  - simul, IOR, mdtest
- Difficult to identify edge-cases or race-conditions observed under real production workloads

The OLCF end-to-end I/O path



Enterprise Disk are connected to controllers via Serial ATA.

20,160
2 TB NL- SAS drives

Enterprise Storage controllers are connected via InfiniBand.

36 DDN SFA12K-40 controller pairs and 16 InifiniBand FDR connections per pair

Lustre servers run parallel file system software and manage incoming FS traffic.

288 Dell servers with 64 GB of RAM each

SION II Network provides connectivity between OLCF resources and primarily carries storage traffic.

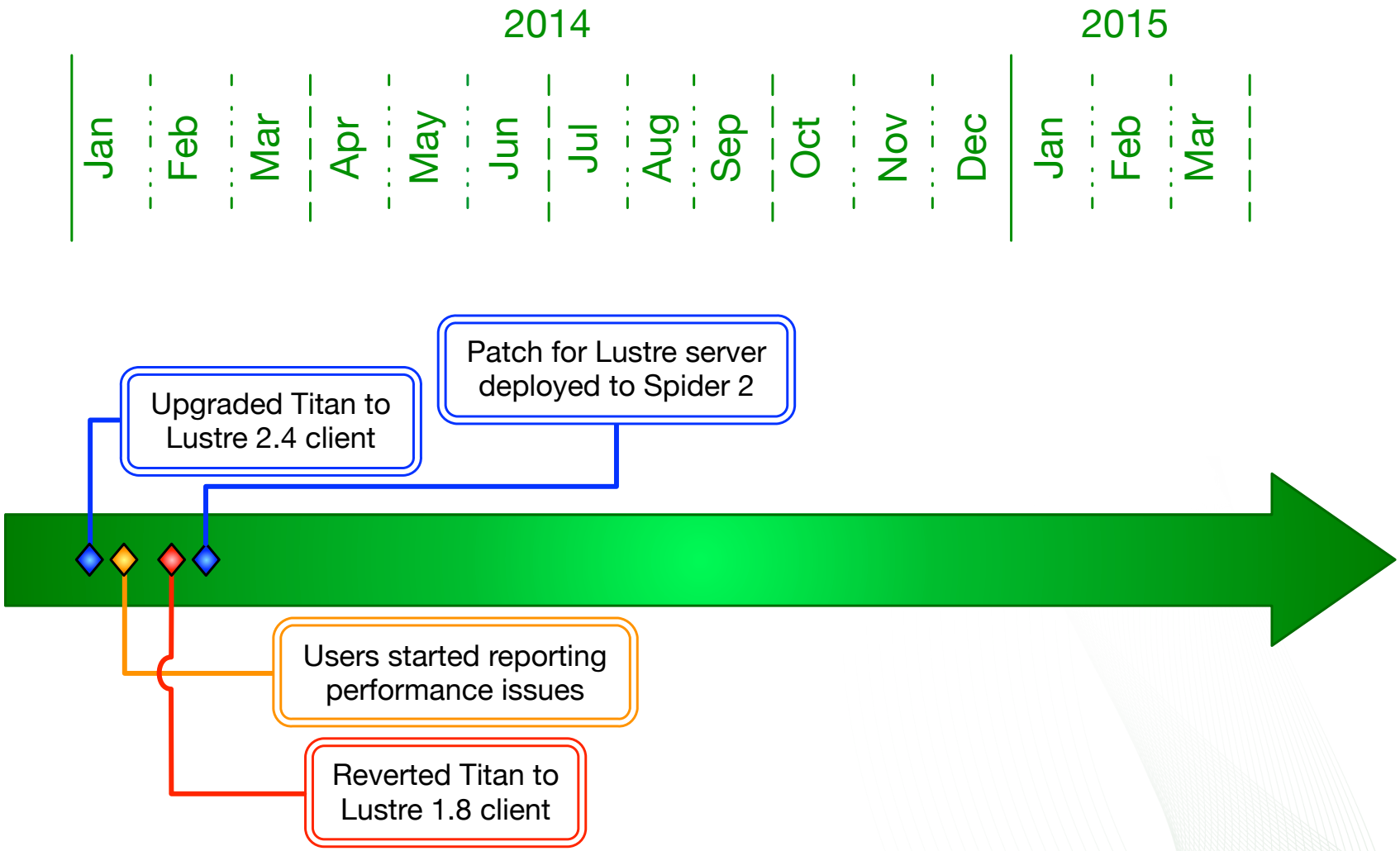
1600 ports, 56 Gbit/sec InfiniBand switch complex

Lustre Routers and Clients run parallel file system client software and forward I/O operations from HPC clients.

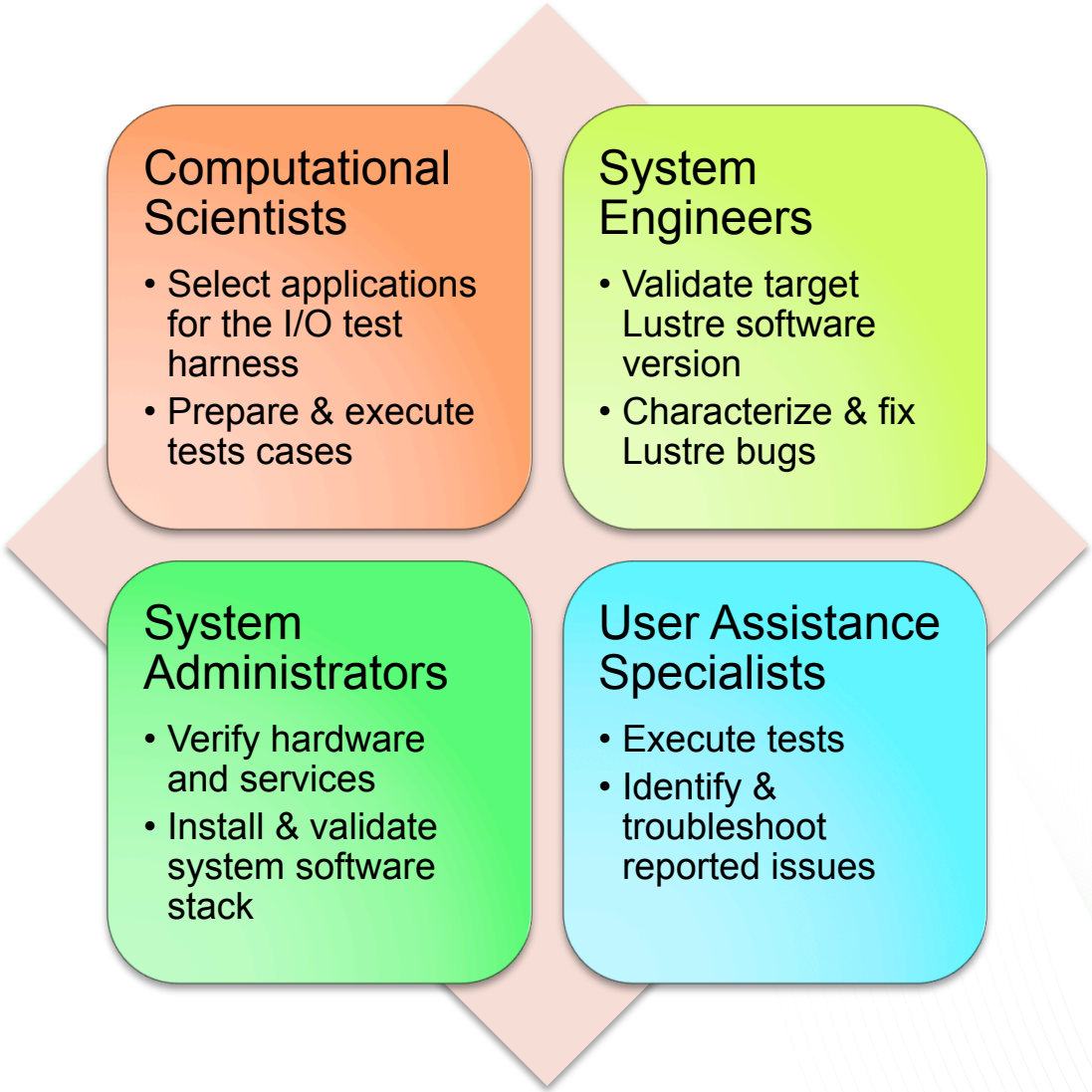
440 XK7 XIO nodes configured as Lustre routers on Titan

SPIDER 2 | SION 2 | COMPUTE

Upgrading Lustre on Spider 2



OLCF File Systems Testing Team



I/O test harness

- User applications at OLCF represent a diverse scientific community
- Broad range of programming models and I/O patterns
  - Single shared file, file-per-process
- Evaluates functionality, stability, and performance
- Includes micro-benchmarks, application I/O kernels, real-world applications
- Executes file system operations users frequently perform
  - ls, find, grep, cp, tar
- Simulates entire users' workflow
  - Application building, job submission, analysis of results
- Infrastructure also used in acceptance

I/O test harness applications

S3D
Combustion modeling in flames
- Fortran I/O
- MPI
- 3 checkpoints
- 3.3MB file
- Stripe count of 1

GTC
Particle-in-cell simulations of plasma microturbulence in fusion
- Fortran I/O
- MPI+OpenMP
- 1 checkpoint
- 274MB file
- Default striping

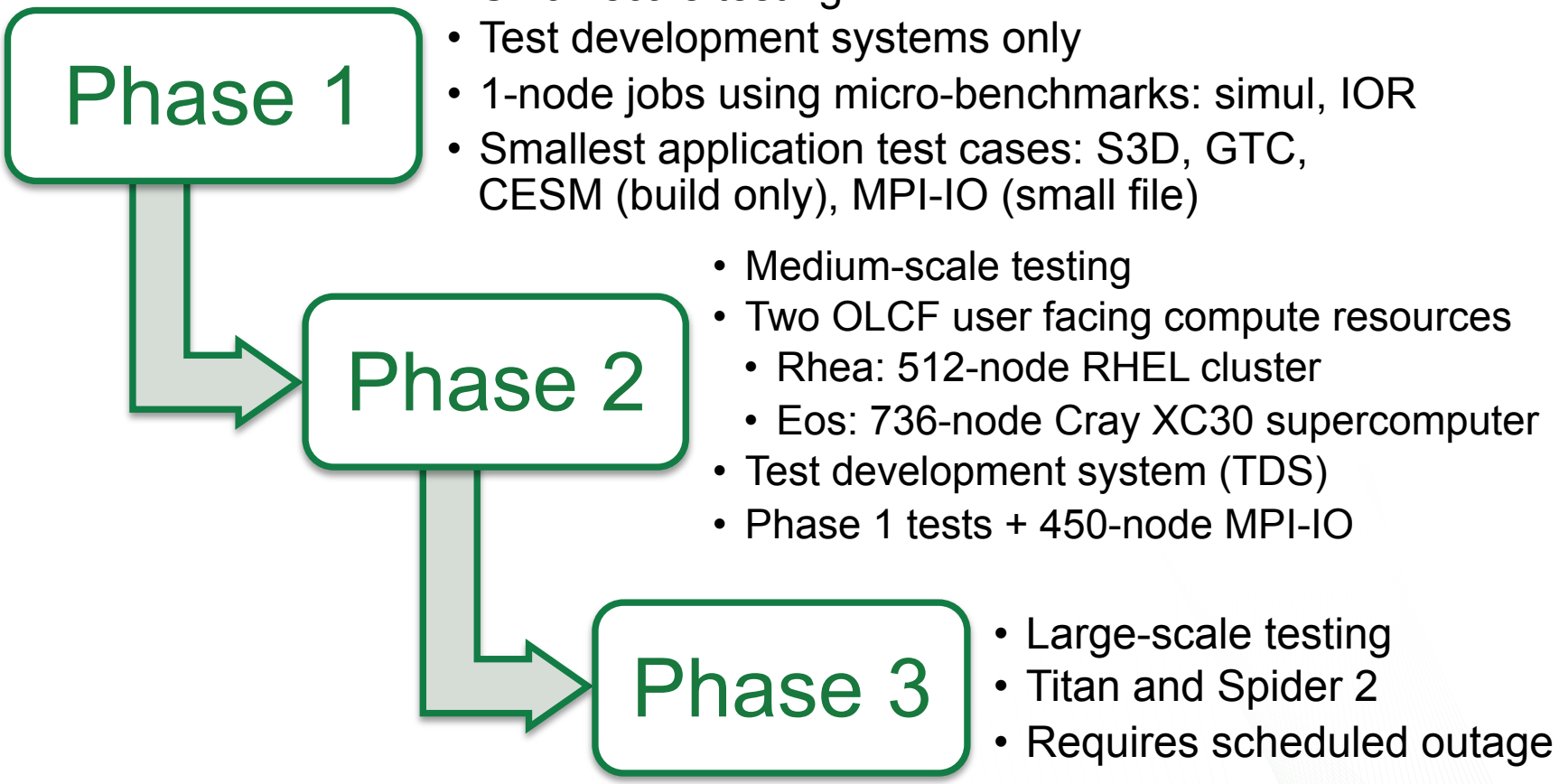
FLASH (I/O kernel)
Multiphysics code used to simulate astrophysical environments
- HDF5 library
- 625 nodes
- Single read/write
- 54GB file
- Default striping

MPI-IO
Large single shared file
- Collective MPI-IO
- 3 successive reads and writes
- 512GB file
- Varying stripe counts
- Increasing number of clients

CESM
Community Earth System Model
- pnetCDF library
- 3,567 nodes
- Record build time
- One test case

User operations
Test cases built from user reports
- Copy large directory tree (100k small files)
- Build applications
- Single client reads with varying block size

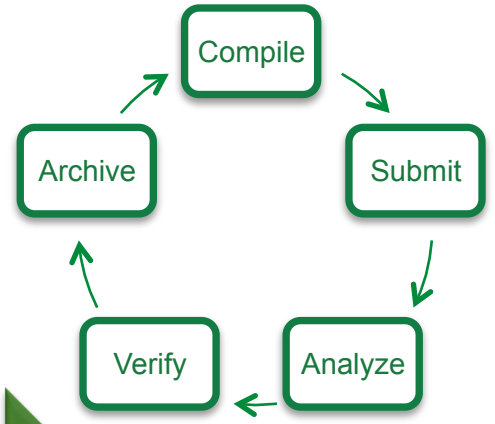
Testing Methodology



Testing on Titan

- Scheduled to coincide with planned maintenance windows
- Short testing window (2-4 hours)
  - Must carefully select test cases
- Goal is to test functionality, stability, and performance
- Divided into stages:
  - initial stability test
  - basic scaling tests
  - hero runs for performance
  - application harness to simulate a realistic workload

Testing on Titan (cont'd)



Basic scaling tests

IOR:

- 1 to 4,096 nodes
- 4 to 1,008 stripes
- Fixed striping/ increasing number of clients
- Fixed number of clients/increasing stripe count

Initial Stability

MPI-IO:

- 450 and 1,024 clients
- 512GB file
- Stripe counts of 32, 128, 360, 720
- Running continuously

Realistic Workload

Full-suite:

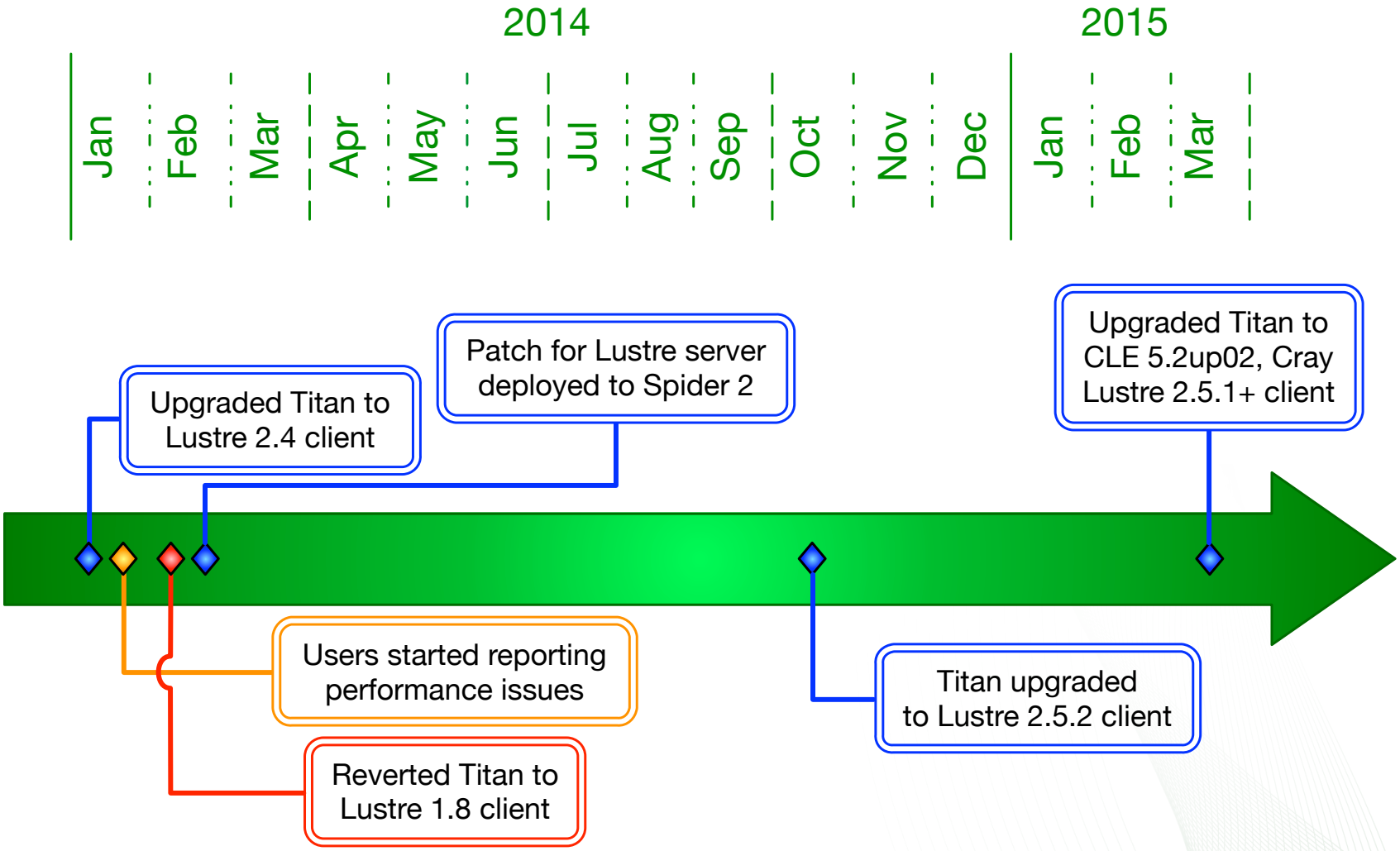
- S3D
- GTC
- CESM (build and run)
- FLASH I/O kernel
- MPI-IO
- User operations

Performance

Hero runs:

- Maximum performance
- Ideal circumstances

Upgrading Lustre on Spider 2 (round 2)



Lustre bugs identified

Date	BugID	Identified By	Description
Feb 26, 2014	LU-4008	Application I/O harness	vmalloc contention on MDS
Mar 20, 2014	LU-4719	Synthetic Benchmarks	Kernel panic with large stripe files
May 8, 2014	LU-4578	Application I/O harness	Adaptive timeout bug causing MDS to reboot
May 8, 2014	LU-4584	Application I/O harness	Client evictions in Lustre 2.4
Jul 4, 2014	LU-5294	Synthetic Benchmarks	Cannot unlink or rm
Aug 7, 2014	LU-4829	Synthetic Benchmarks	Crash on mount
Oct 14, 2014	LU-5803	Application I/O harness	Recovery issues, server not able to keep up with requests
Oct 16, 2014	LU-5724	Application I/O harness	Imperative recovery issues

Summary

- A file system should be tested for functionality, correctness, and performance throughout its lifetime
  - Not only initial acceptance
- Synthetic benchmarks alone cannot adequately stress the end-to-end I/O path
- The I/O test harness is a combination of synthetic benchmarks, application kernels, and real-world scientific codes
- Testing at increasing scales minimizes user impact
  - Outages on production systems are only scheduled if small and medium-scale tests are successful
- Testing methodology also used for system software upgrades

Future Work

- Periodically evaluate the workload on Titan to update the I/O test harness
- Extend test cases to include additional I/O libraries
- Continue adding reproducers for user reported edge-cases

Acknowledgements

- The entire OLCF file systems testing team, which also includes: Don Maxwell, Matt Ezell, John Lewis, and Phil Curtis.
- Duane Rosenberg, Ramanan Sankaran, Wayne Joubert, Matt Norman, and Bronson Messer for providing applications for the test suite.
- Arnold Tharrington for use of his harness scripts to automate the build, launch and analysis of tests.
- This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Questions?

Thank you!