Technical Overview of the OLCF's Next Generation Parallel File System

Galen Shipman, David Dillow, Douglas Fuller, Doug Reitz, Raghul Gunasekaran, Jason Hill, Youngjae Kim, Sarp Oral, James Simmons, Feiyi Wang

Presented by: Douglas Fuller and Jason Hill

Agenda

- How did we get here
- Operational successes
- Gemini experiences
- Workload characterization
- Procurement and benchmarks
- Next generation architecture
- Questions

Transitioning I/O to next gen computing

- From Jaguar to Titan
  - Number of cores: 224K → 300K
  - Memory: 300 TB → 600 TB
  - Peak Performance: 2.2 PFlops → 10-20 Pflops
  - Proprietary Interconnect: SeaStar2+ → Gemini
  - Peak egress I/O (over IB): (192 x 1.5 GB/s) → (384-420 x 2.8-3 GB/s)

More capable platform for science → more demanding I/O requirements to deliver the science

Starting from Spider …

- Spider → Next gen parallel file system
- Designing, deploying, and maintaining Spider was a trail blazer
  - No ready available solution at the time of design or deployment
  - Novel architecture
- Center-wide shared file system approach
  - Eliminating islands of data
  - Decoupled file system from compute and analysis platforms
  - Rolling or partial upgrades possible with no down time
  - *Single-point of failure?*

Spider availability

- Scheduled Availability (SA)
  - % of time a designated level of resource is available to users, excluding scheduled downtime for maintenance and upgrades

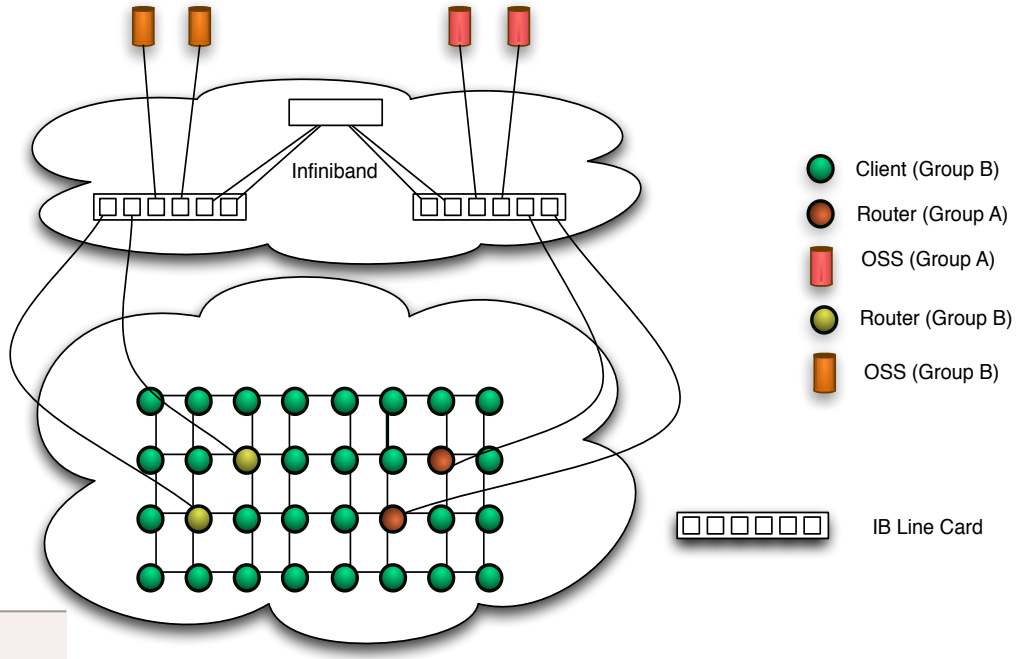
System	Scheduled Availability (SA)			
	2010 Target	2010 Actual	2011 Target	2011 Actual
Widow1	95.0%	99.7%	95.0%	99.26%
Widow2	NIP	NIP	95.0%	99.34%
Widow3	NIP	NIP	95.0%	99.36%

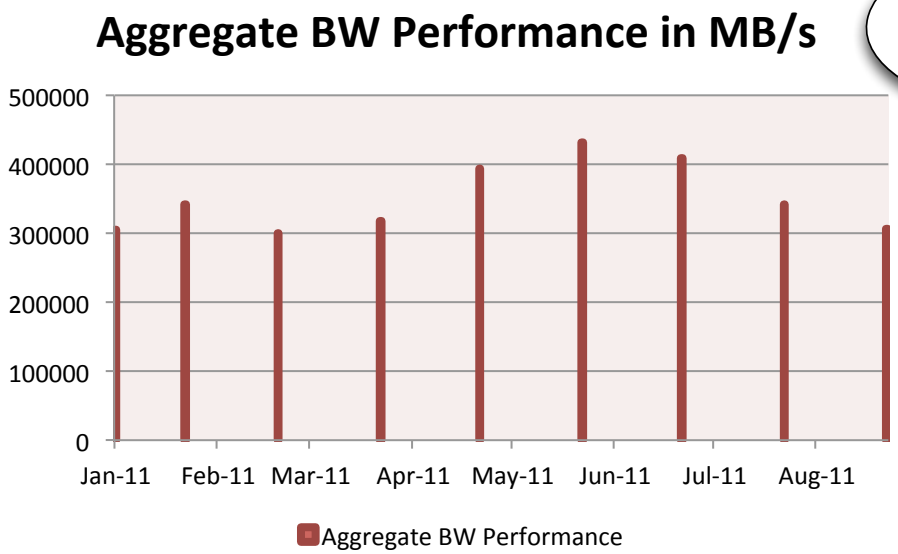
- Widow1
  - **100% availability in 8 of the 12 months of 2011** with SA of 99.26% over the entire year
- Availability and reliability surpassed our expectations

Next gen file system will also be center-wide shared architecture

Advanced LNET Routing

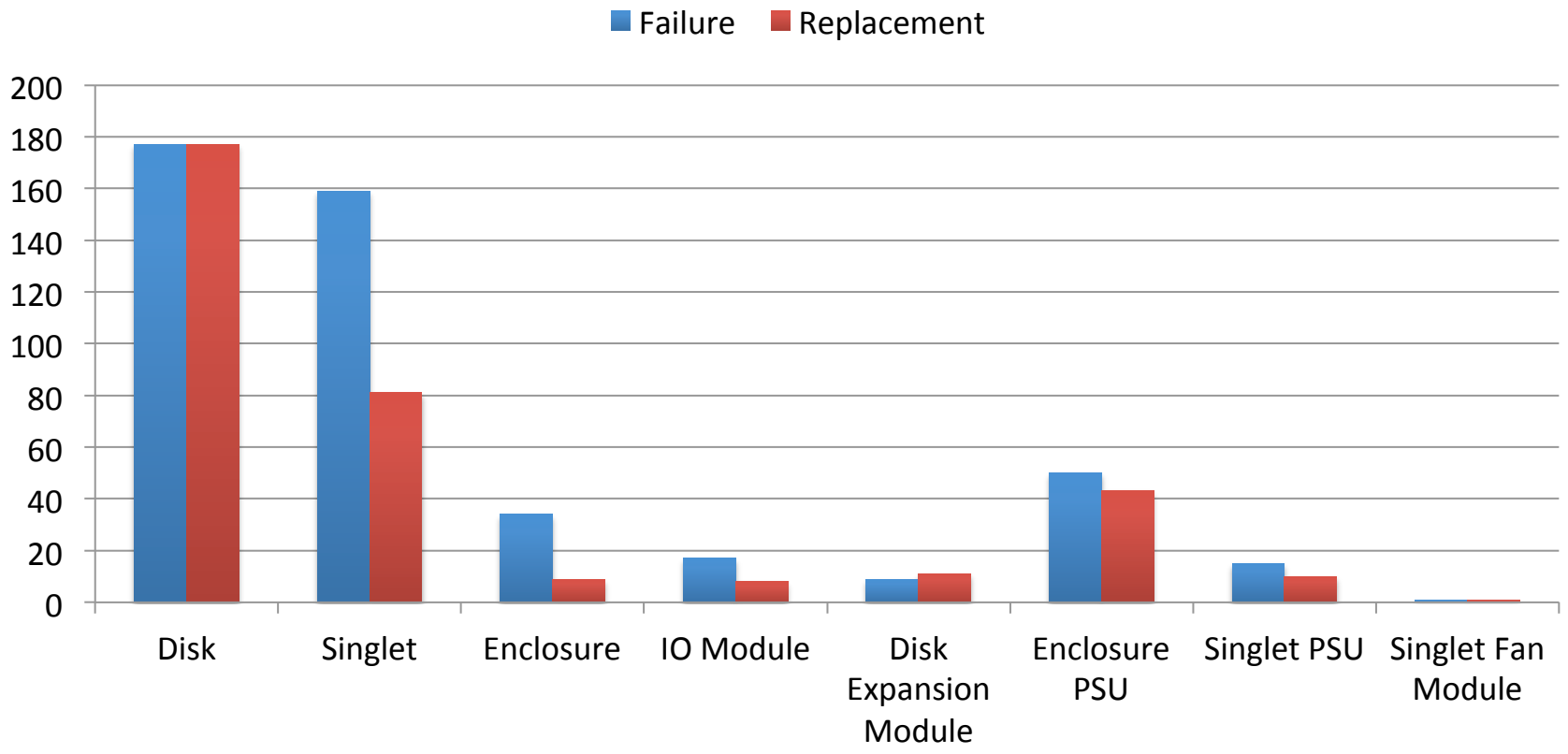
- Implemented Fine-Grained Routing in May 2011
- Application perf increases
  - Up to 37% for writes
  - Up to 15% for reads
- Had to be removed for upgrades in August





DDN 9900 Reliability

- Distinguish between cases and actual component failure requiring replacement
- Because of architecture most are **not** service interruptions

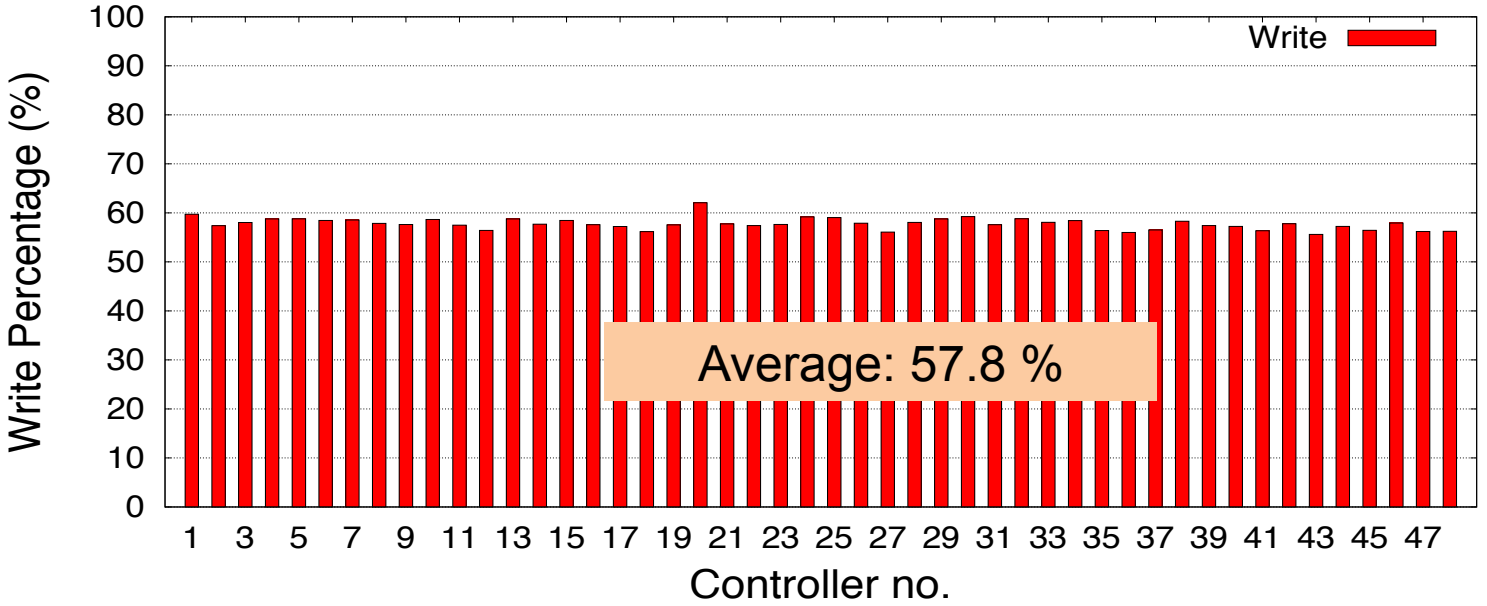


SeaStar2+ to Gemini

- Increased reliability/stability
- 793 GB/s from 370 XIO nodes for LNET Self Test
- One outstanding issue being addressed in short term
- Should provide enough bandwidth for our target
  - Tweaks to buffering/credits; improvements in LNET checksumming?
- Early on a lack of documentation

I/O Workload Characterization

- "Workload characterization of a leadership class storage cluster"
  - http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5668066



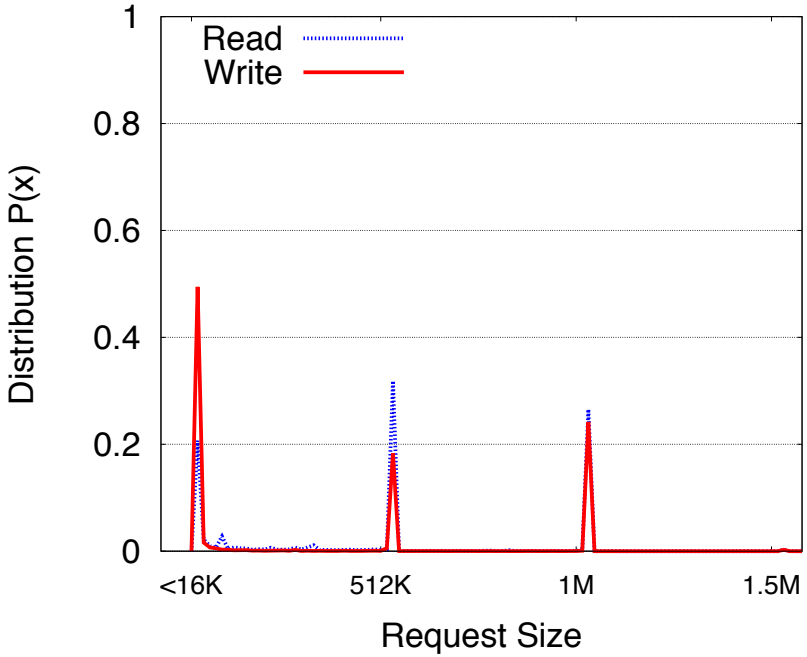
42.2% Read requests ➔ still significantly high!!!

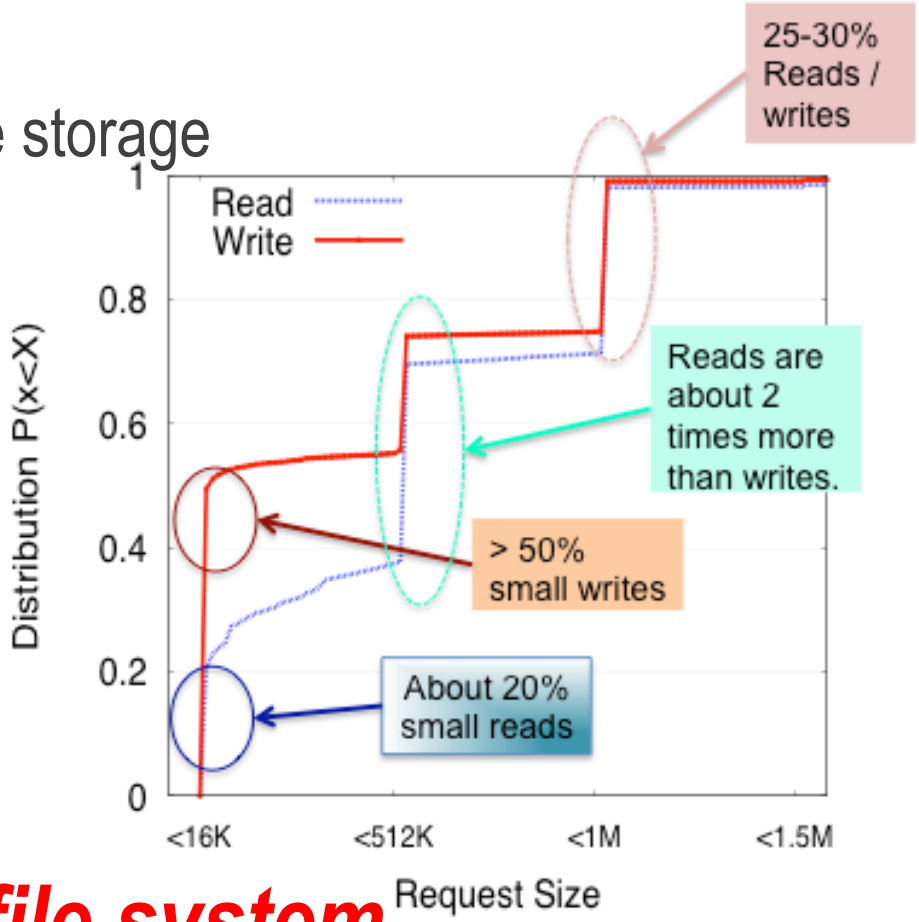
Next gen file system

— ***can not only be optimized for checkpointing***

I/O Workload Characterization

- Many requests are small (< 16 KB), while majority are (< 16 KB, 512 KB, or 1 MB)
- Most I/O is random once it hits the storage



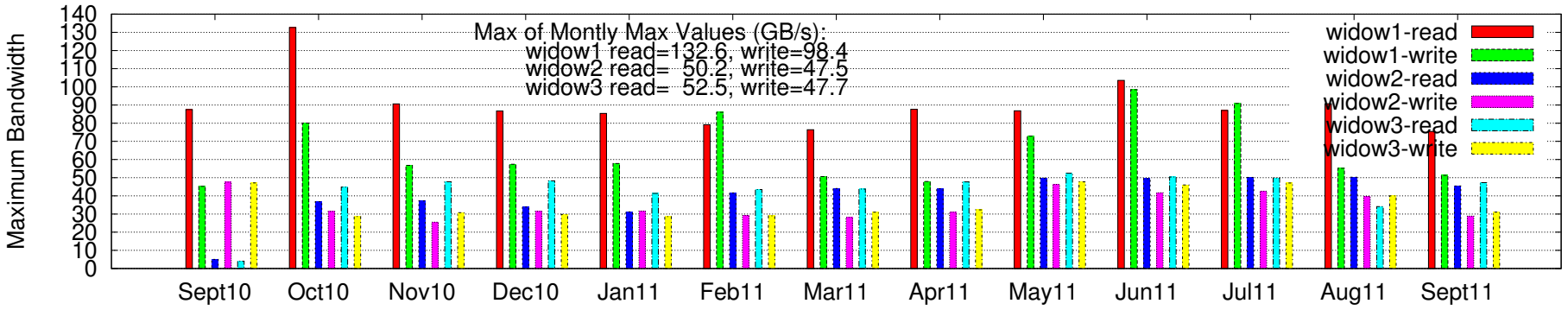


Next gen file system

—should support mixed workloads

I/O Workload Characterization

- Capturing maximum bandwidth per filesystem
  - As seen at the storage controller level
- Captures both day-to-day workload and testing periods
- Examine on a monthly basis
  - Could do higher frequency



Procurement

- Acquisition process
  - Open procurement
  - Timetable: TBD (2012-2013 timeframe)
- Procurement benchmarks
  - Publicly available
    - http://www.olcf.ornl.gov/wp-content/uploads/2010/03/olcf3-benchmark-suite.tar.gz

Procurement Benchmarks

- Based on lessons learned from the workload characterization work
- Targeted to evaluate offered solutions and new technologies for the procurement
- Complete package to run tests, gather and parse data, and to plot summary results
- Block I/O level benchmarks
- Lustre file system level benchmarks

Procurement Benchmarks

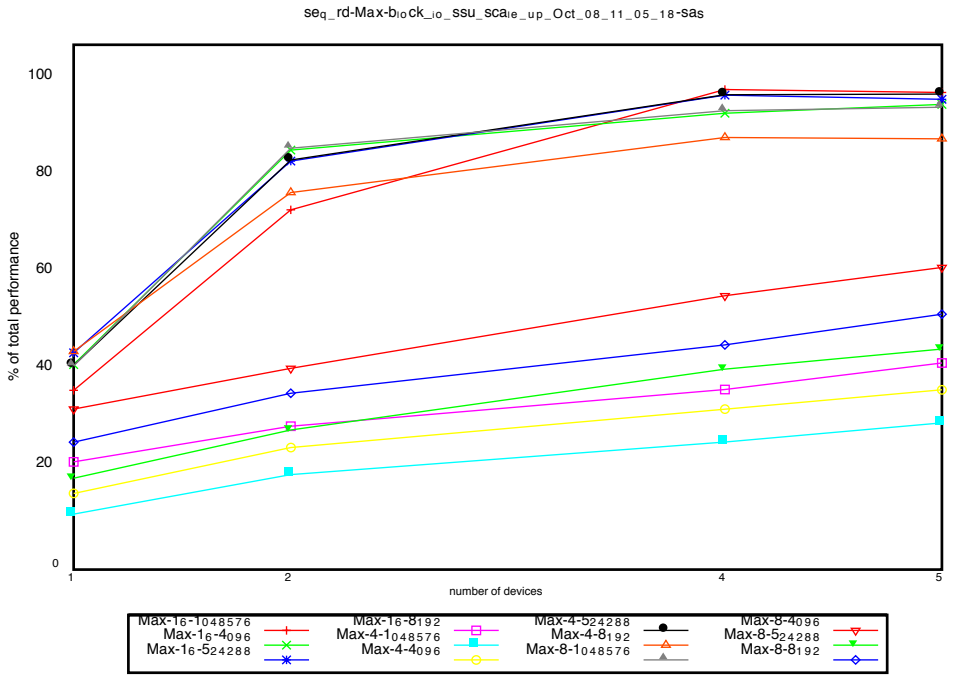
- Block I/O
  - Libaio based, fair-lio as I/O engine
  - Single host single LUN
    - Profile single target; single host
    - 720 tests, total run time 9 hours (720 tests x 45 seconds)
  - Single host all LUNs
    - Profile single host, multiple targets
    - Number of tests varies with number of targets
    - $\log_2$(number of targets) x (144 tests x 45 seconds), 7.2 hours for 5 targets

Procurement Benchmarks

- Block I/O
  - SSU all LUNs – healthy
    - Profile an SSU in a healthy state
    - Number of tests varies with number of targets
    - $\log_2$(number of targets) x (144 tests x 45 seconds), 7.2 hours for 5 targets
  - SSU all LUNs – degraded (10% targets in rebuild)
    - Profile an SSU in a degraded state
    - Identical number of tests and setup to the SSU healthy mode tests

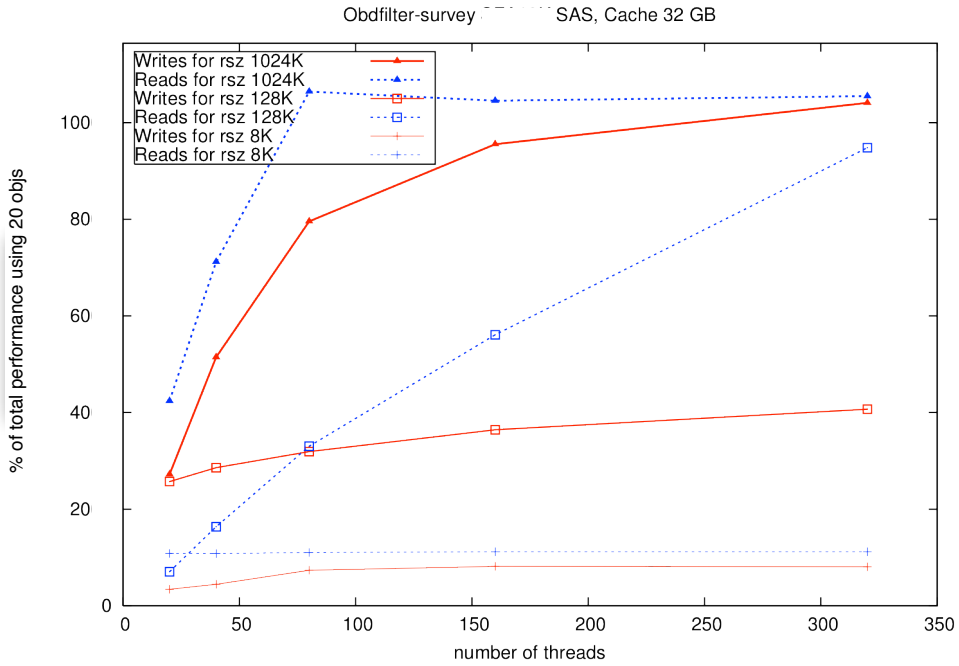
Procurement Benchmarks

- Block I/O
  - A list of test parameters are generated a priori to the execution
    - I/O mode (sequential or random), operation (read or write)
    - Number of targets to exercise (not valid for the single host single LUN test)
    - Queue depth, block size, number of iterations
  - Generated list of test parameters are randomized a priori to the execution
  - Randomized list of test parameters are fed into the fair-lio engine
    - Eliminating cache effects on the host side as well as the controller side



Procurement Benchmarks

- Lustre file system
  - Obdfilter-survey based
  - OLCF wrapper to generate certain test conditions and feed them to the obdfilter-survey
  - OSS level testing, no clients are needed
  - Development and execution tested with Lustre version 1.8



New Architecture

- Target numbers for next gen parallel file system
  - 1 TB/s file system-level well-formed I/O performance
  - 240 GB/s file system-level random I/O performance
  - Capacity will be based on the selected storage media
    - Expected to be 9-20 PB
  - Availability: >95%
    - Expected availability will be similar of Spider's

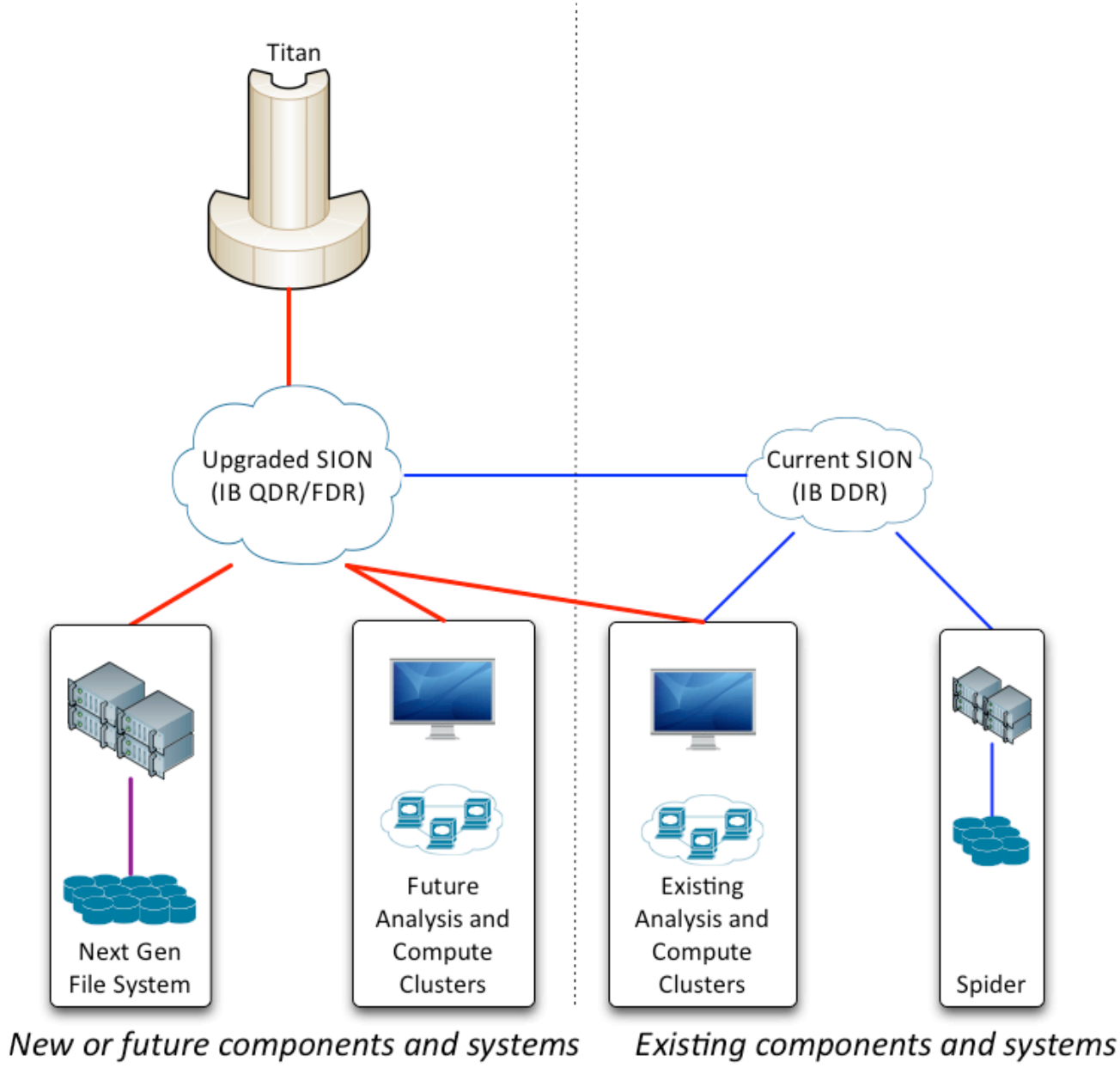
Architecture

- Expected storage and network architecture
  - Will be built using scalable building blocks (SSU)
  - Host-side connectivity: IB FDR or QDR
    - SION tech refresh and upgrade
  - Disk-side connectivity: FC, IB, SAS, …
    - Agnostic of the host-side

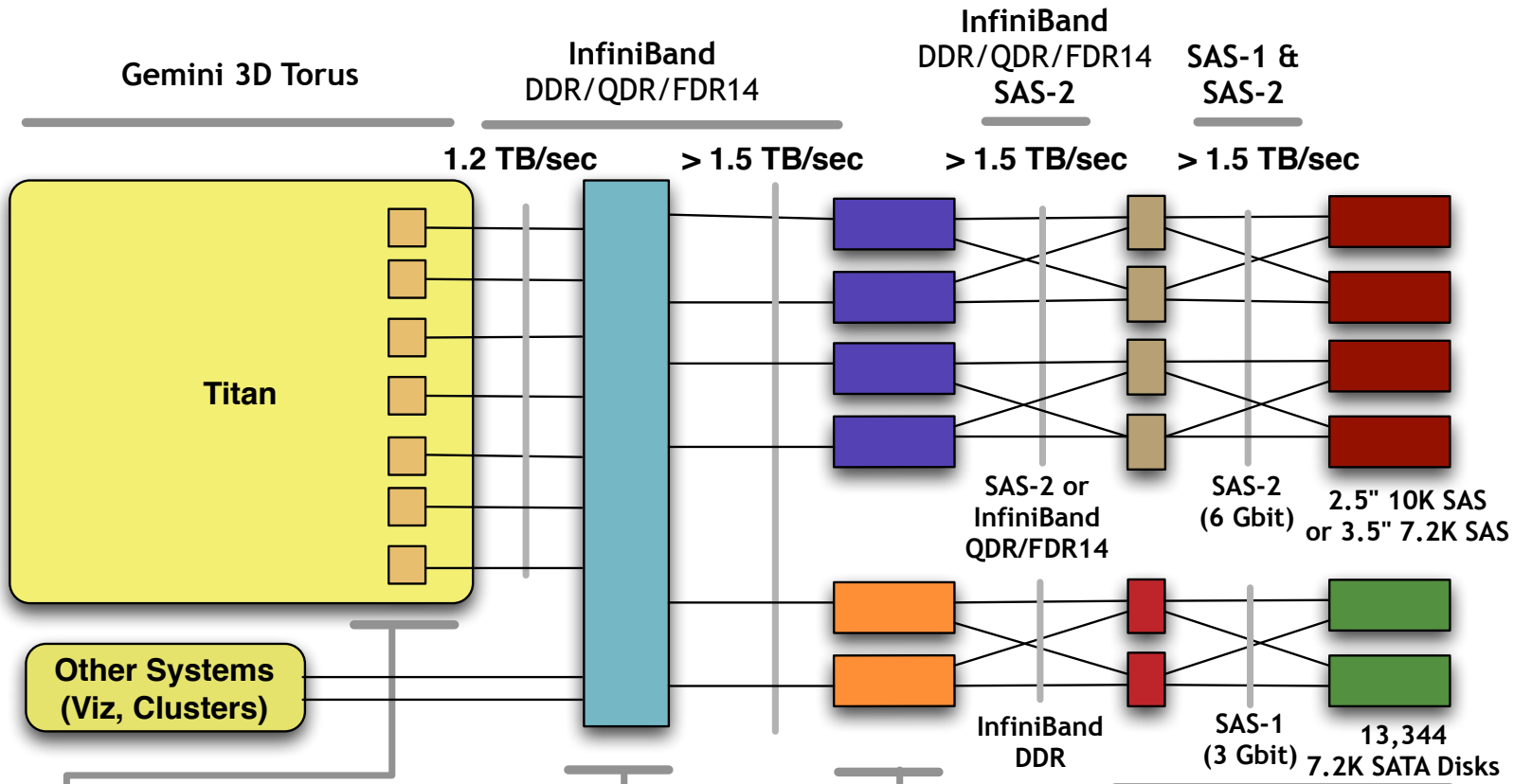
Another advantage of decoupled parallel file system architectures

  - Next gen file system and Spider will be online concurrently
    - Spider will be connected to the upgraded SION through existing SION
    - Spider EOL expected to be 2014

Architecture



Spider: the Next Generation



Lustre Router Nodes
Forward I/O operations from Compute Clients to Storage Nodes
384 "XIO nodes"

InfiniBand
switch complex
1200+ port 32 or 54 Gbit/sec
3000+ port 20 Gbit/sec InfiniBand switch complex

Storage Nodes
run parallel file system software and manage incoming FS traffic.

Enterprise Storage
controllers and large racks of disks are connected via SAS

Lustre for next gen parallel file system

- Lustre v. 2.2 or later will be used
  - Improved metadata performance
    - pDirOps (2.2)
    - Async glimpse lock (statahead issue)
    - *DNE and SMP scaling*
  - Scalability improvements (2.2)
    - Imperative recovery
    - Wide-striping
    - Portals RPC thread pool
    - *NRS*

Scheduled down-times can be used to harden 2.2 and test future Lustre features, bug fixes, and improvements

Questions?

The research and activities described in this presentation were performed using the resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC0500OR22725.
We gratefully acknowledge the support of the Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under contract DE-AC05-00OR22725 with Oak Ridge National Laboratory.

OAK RIDGE National Laboratory