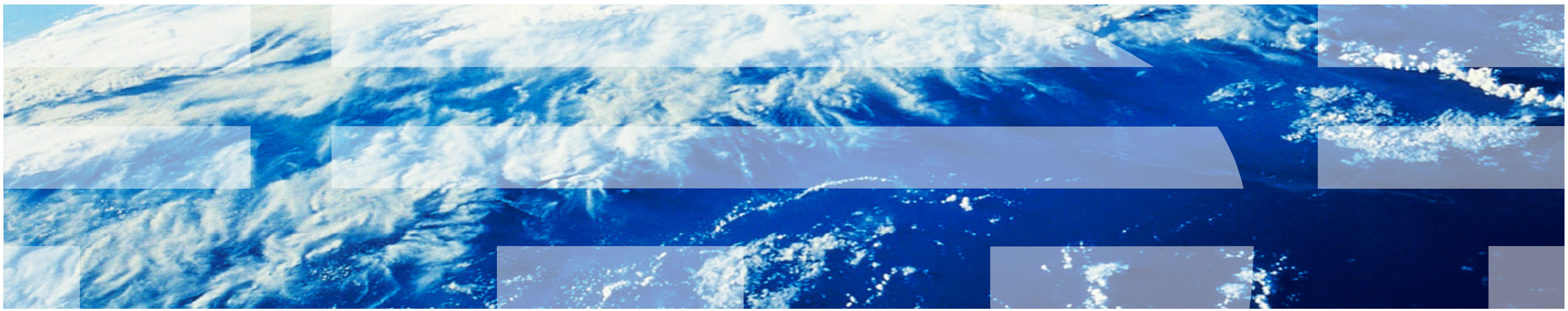
# Storage performance modeling for future systems

Yoonho Park

May 3, 2016



# Agenda

- Storage challenges
- Burst buffer
- APEX workflows document
- Machines analyzed
- LANL workflows performance modeling
- Workflow time distribution
- CORAL burst buffer
- Ongoing work

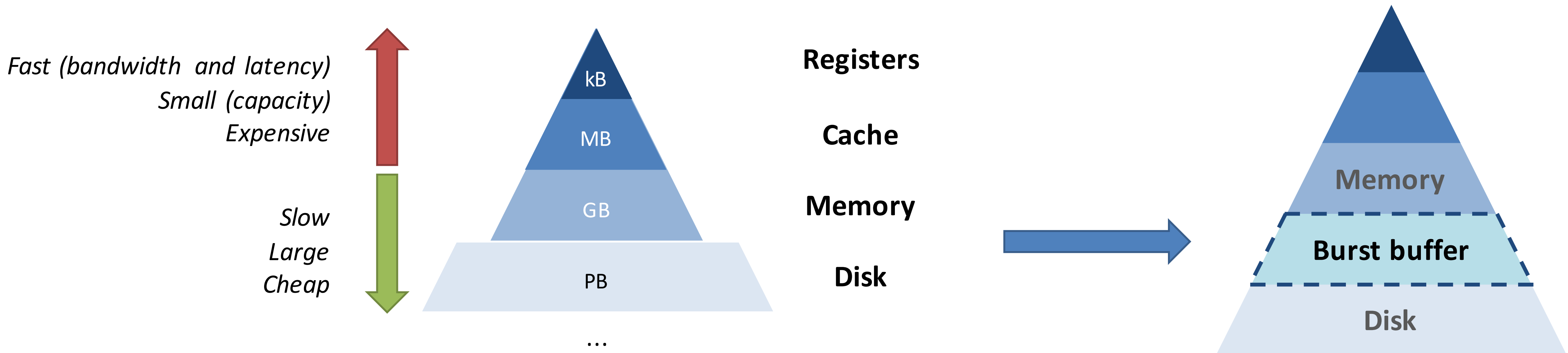
# Storage challenges

- Current parallel file systems are unable to **consistently** deliver an adequate fraction of aggregate disk bandwidth
- I/O patterns that lead to irregularity and unpredictability
  - Multiple processes writing to a shared file (N:1)
  - **Bursty** I/O (e.g. checkpointing) vs **Underutilization** (very low baseline)
- Increased capacity and bandwidth **requirements** for future systems (exascale)

# Burst buffer

- Absorbs bursty I/O patterns via higher bandwidth and lower latency (compared to parallel file system)
- Allows parallel file system to be sized for capacity (not overdesigned)
  - HDD capacity grows faster than bandwidth
  - SSD still is more expensive than disk for capacity

- Use cases
  - Checkpoint and resilience
  - Analysis, post-processing, and visualization
  - Caching and performance optimization
  - Extend memory capacity (e.g. large problems)



# APEX workflows document

- Specification of large-scale scientific simulation and data-intensive workflows
  - Workflow phases
  - Campaign duration
  - Workload percentage
  - Wall time (pipeline duration)
  - Resources allocation (e.g. CPU cores and total memory for routine vs hero runs)
  - Anticipated increase factors (problem size and number of pipelines) by 2020
  - I/O details (e.g. files accessed)
  - Amount of data retained (temporary, campaign, and forever)

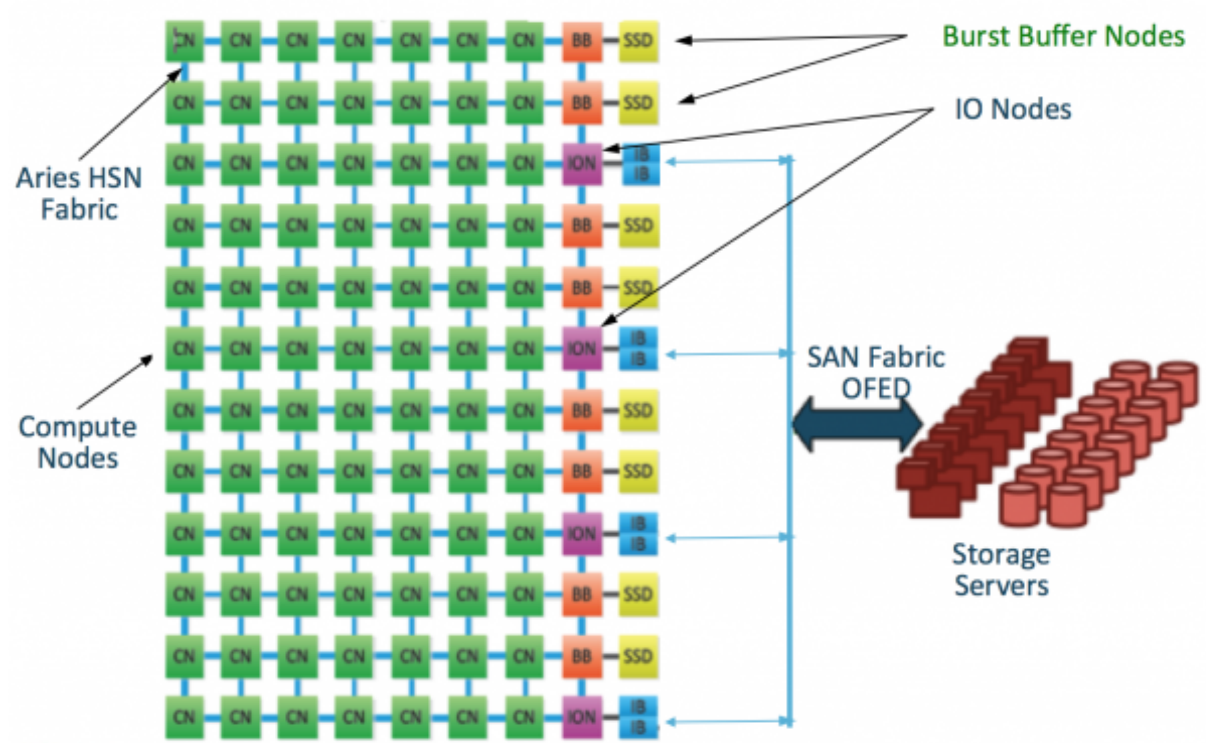
- The information obtained from the document and discussed throughout the meetings with the APEX labs has been used to:
  - **Model** performance improvement provided by having burst buffer for a variety of use cases
  - **Design** and enhance future storage hierarchy architectures and underlying components (e.g. OS support, transparency, and usability)

# Machines analyzed

| | Cielo | Edison | Trinity* |
|---|---|---|---|
| Nodes | 8,944 | 5,576 | 9,436 |
| Total cores | 143,104 | 133,824 | 301,952 |
| Cores per node | 16 | 24 | 32 |
| Total memory (TB) | 286 | 357 | 1,208 |
| Memory per node (GB) | 32 | 64 | 128 |
| Bandwidth per node (GB/s) | 85 | 103 | 137 |
| PFS capacity (TB) | 7,600 | 7,560 | 82,000 |
| BB capacity (TB) | - | - | 3,700 |
| PFS bandwidth (TB/s) | 0.16 | 0.17 | 1.45 |
| BB bandwidth (TB/s) | - | - | 3.30 |

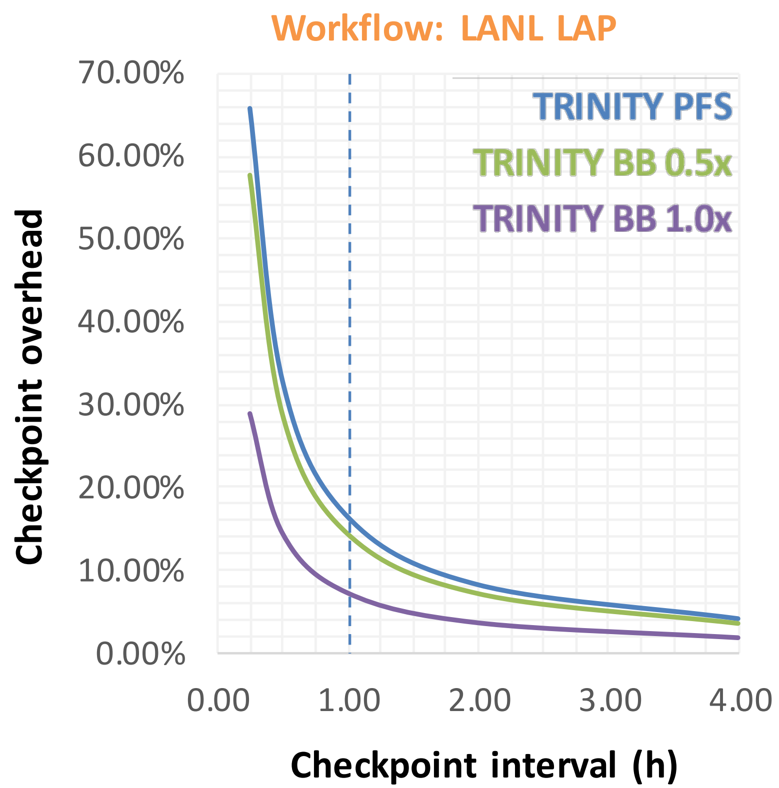
* CPU only (no accelerators)
Trinity data obtained from [4]

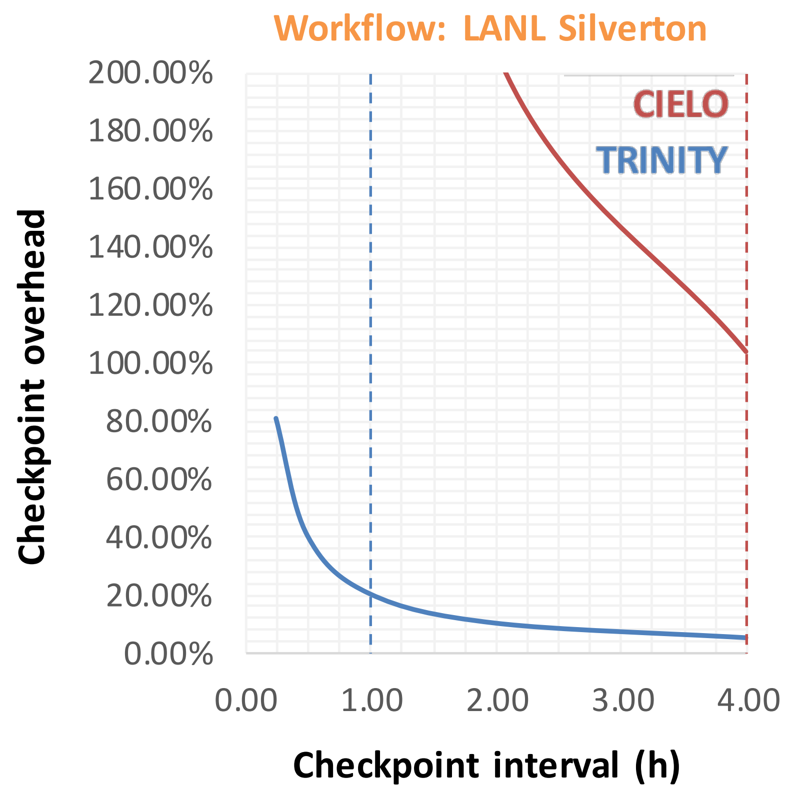
- Trinity burst buffer to main memory ratio: 1.75 X
- Application efficiency estimated to be 88% (12% of checkpoint overhead) [3]
- Trinity burst buffer nodes:



# LANL workflows performance modeling

- Performance modeling (predictions) based on anticipated increased problem sizes for 2020
- Around 2x I/O performance improvement from parallel file system to burst buffer
  - Graph also shows results for half BB bandwidth
- Around 20x improvement over Cielo parallel file system (for the same checkpoint interval)
  - Essential to maintain checkpointing feasible



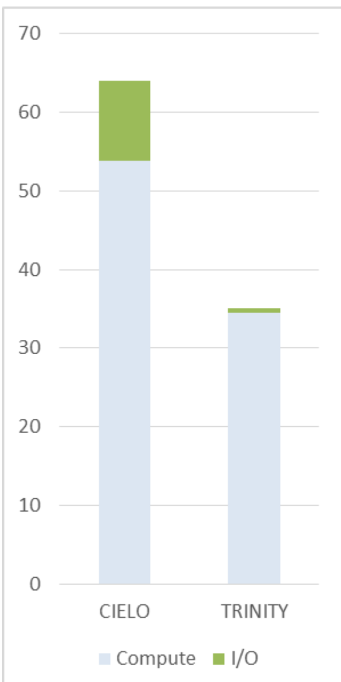



# Workflow time distribution

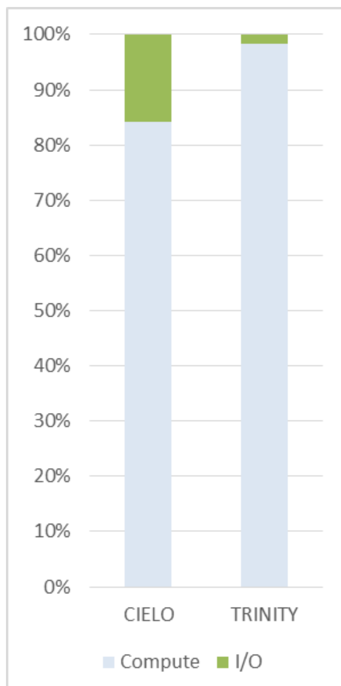
- Drastic performance improvement for checkpointing and other I/O operations

Predictions based on checkpoint interval of 1 hour and current problem size (hero run, without increasing factors)

**Workflow: LANL LAP**

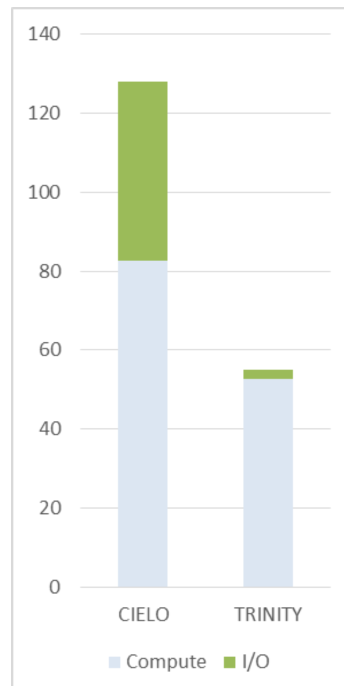



**hours**

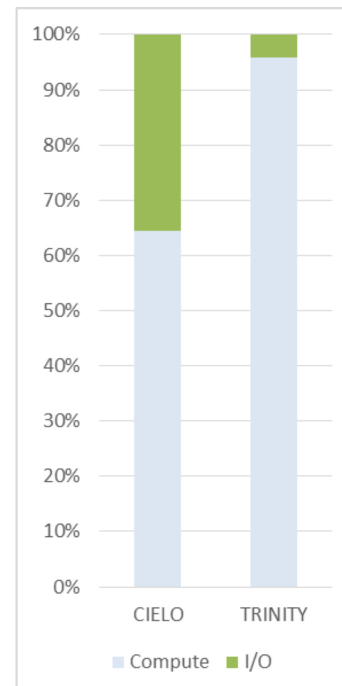



**percentage**

**Workflow: LANL Silverton**




**hours**




**percentage**

# CORAL burst buffer

- Support rapid checkpoint/restart to reduce the parallel file system performance requirements by an order of magnitude (bandwidth)
- **Asynchronous drain** checkpoint data to CORAL parallel file system
- **Per-node design** to maximize throughput and minimize latency to utilize the burst buffer for checkpointing

- **Deterministic** performance
  - Burst buffer bandwidth variation should not exceed 5% and must not degrade over a period of 5 years
- **Reliability** of the burst buffer is a function of node electronics and SSD drive
  - MTTF of more than 2 million hours
  - Mean time to data loss solely based on SSD is designed to be at least 434 hours (4,608 nodes)
  - Burst buffer is non-volatile, data can still be retrieved up to three months after node failure or power outage

# Ongoing work

- Workflows specification is also being used to model other performance characteristics (e.g. processing, memory, and networking)
- Modeling performance, cost, and other aspects of different burst buffer architectures (e.g. per-node vs specialized burst buffer nodes)