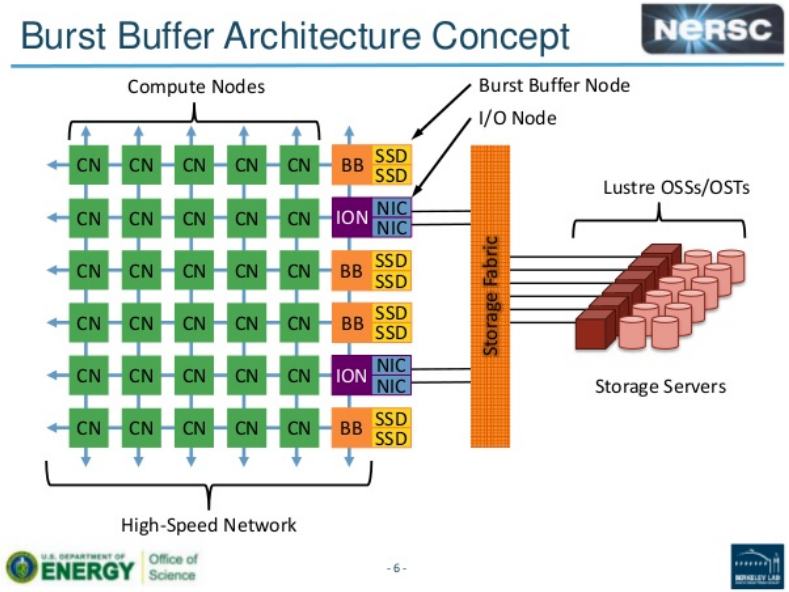
Evaluating Performance of Burst Buffer Models for Real-Application Workloads in HPC Systems

Harsh Khetawat
Frank Mueller
Christopher Zimmer

Introduction

- Existing storage systems becoming bottleneck
- Solution: burst buffers
- Use burst buffers for:
  - Checkpoint/Restart I/O
  - Staging
  - Write-through cache for parallel FS



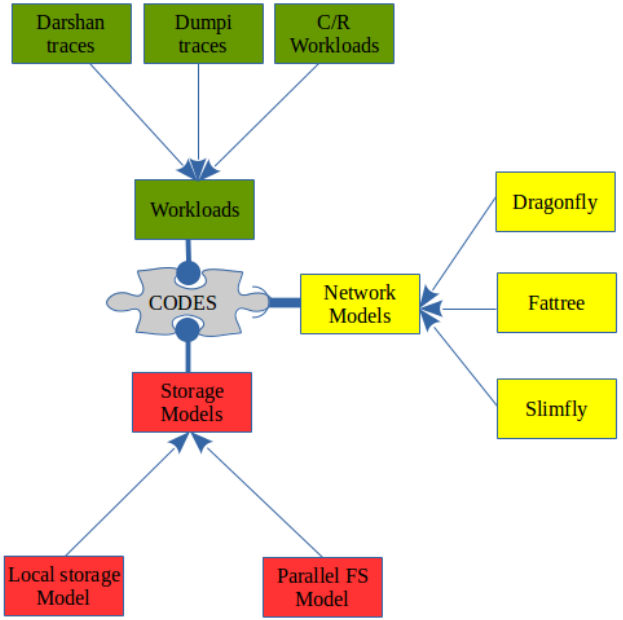
Burst Buffers on Cori

Placement

- Burst buffer placement:
  - Co-located with compute nodes (Summit)
  - Co-located with I/O nodes (Cori)
  - Separate set of nodes
- Trade-offs in choice of placements
  - Capability – I/O models, staging, etc.
  - Predictability – Impact on shared resources, runtime variability
  - Economic – Infrastructure reuse, cost of storage device
- I/O performance dependent on placement
  - Choice of network topology

Idea

- Simulate network and burst buffer architectures
  - CODES simulation suite
  - Real-world I/O traces (Darshan)
  - Full multi-tenant system with mixed workloads (capability/capacity)
  - Supports network topologies
  - Local & external storage models
- Combine network topologies and storage architectures
- Performance under striping/protection schemes
- Reproducible tool for HPC centers



Conclusion

- Determine based on workload characteristics:
  - Burst buffer placement
  - Network topology
  - Performance of striping across burst buffers
  - Overhead of resilience schemes
- Reproducible tool to:
  - Simulate specific workloads
  - Determine best fit

Thank You