Virtualization-based Bandwidth Management for Parallel Storage Systems

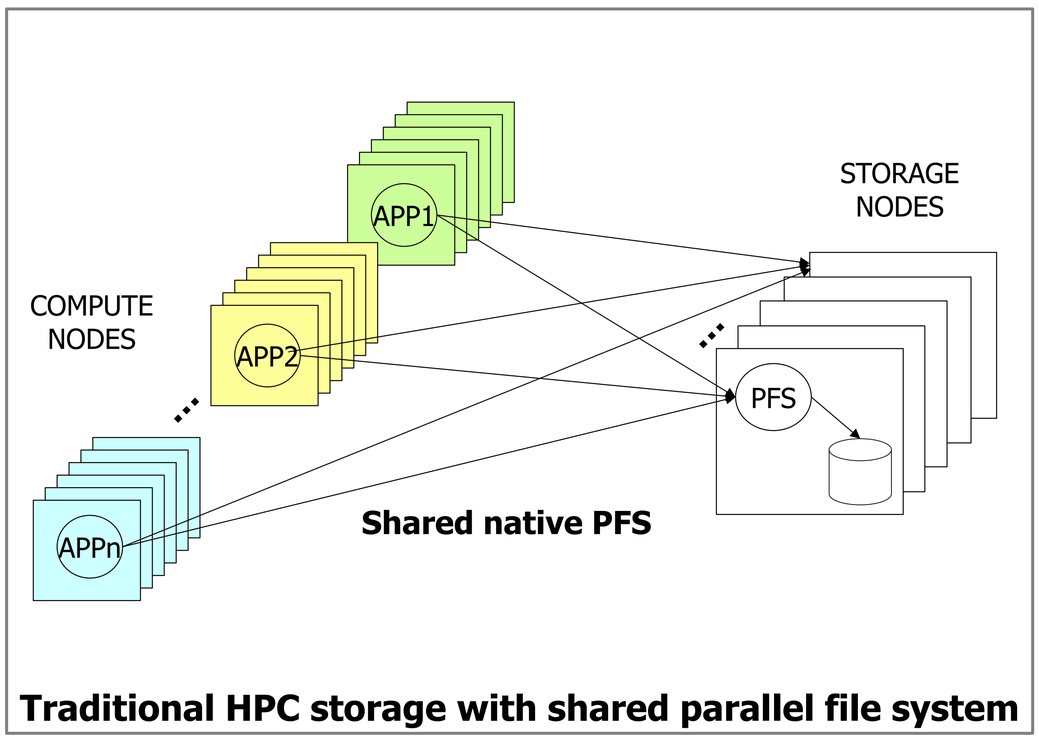
Yiqi Xu, Lixi Wang, Dulcardo Arteaga, Ming Zhao *School of Computing and Information Sciences, Florida International University*
Yonggang Liu, Renato Figueiredo *Electrical and Computer Engineering, University of Florida*

Overview

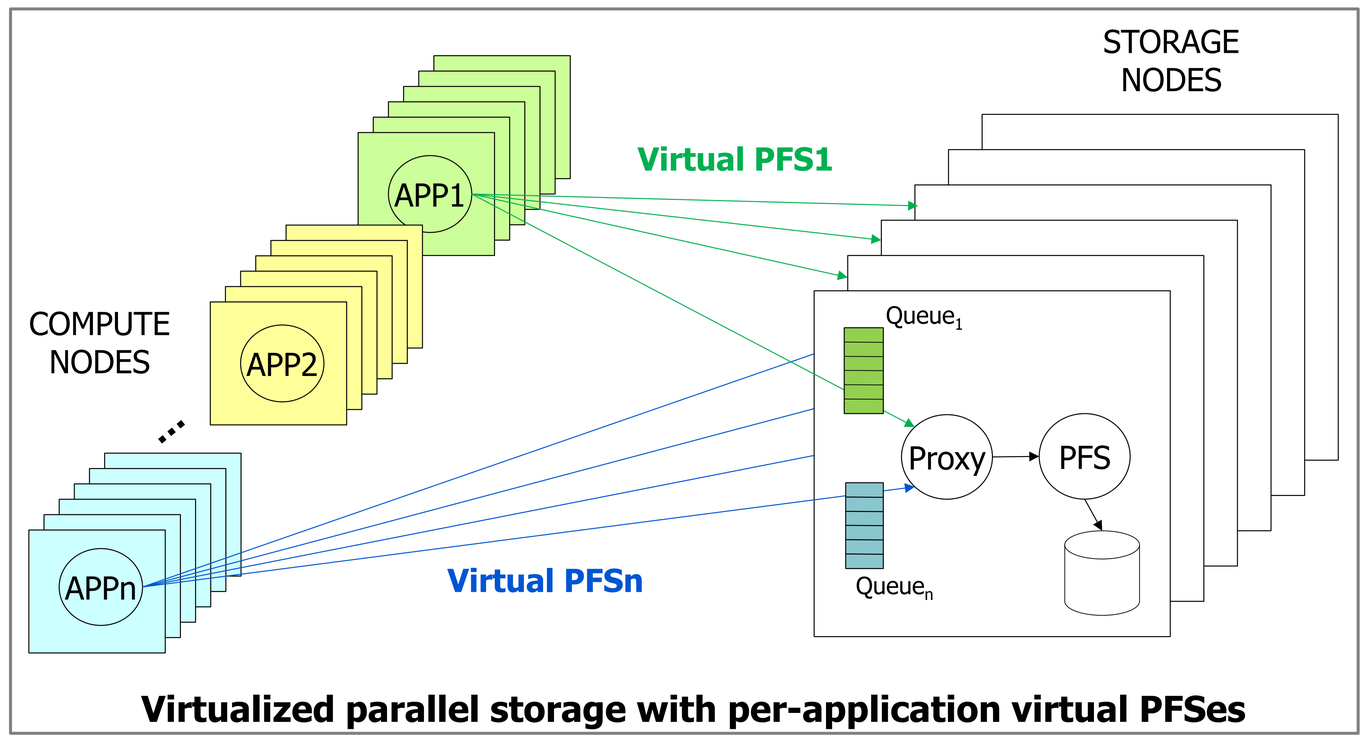
- **Goal:** Application Quality of Service (QoS) driven parallel storage bandwidth management
- **Challenges:**
  - The lack of QoS differentiation in typical high-performance computing (HPC) storage systems
  - The diversity in HPC applications' I/O access patterns and requirements
- **Solution:** Parallel file system (PFS) virtualization based storage management

Parallel File System Virtualization based Bandwidth Management

- **Parallel File System Virtualization**
  - Enable per-application virtual PFSs upon shared physical PFS deployment (e.g., PVFS2, Lustre, GPFS, PanFS, etc.)
  - Allow virtual PFSs to be dynamically created and destroyed based on application lifecycles
  - Allocate parallel storage bandwidth across virtual PFSes according to the applications' I/O access patterns and requirements
- **Proxy-based PFS virtualization**
  - Interpose between native PFS clients and servers to broker the parallel I/Os
  - Create per-application virtual PFSes by identifying different applications' I/Os and isolating them with independent queues
  - Enforce bandwidth allocation by scheduling I/Os across per-application queues according to the sharing algorithm and policy



Traditional HPC storage with shared parallel file system

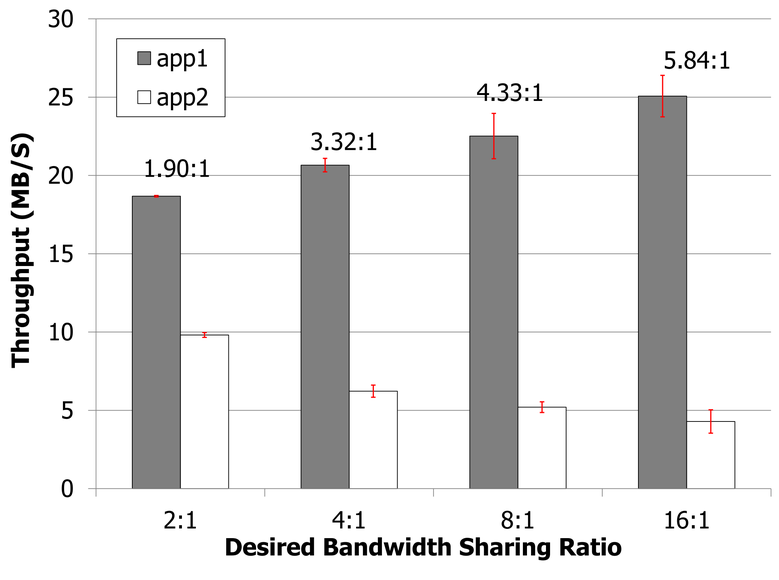


Virtualized parallel storage with per-application virtual PFSes

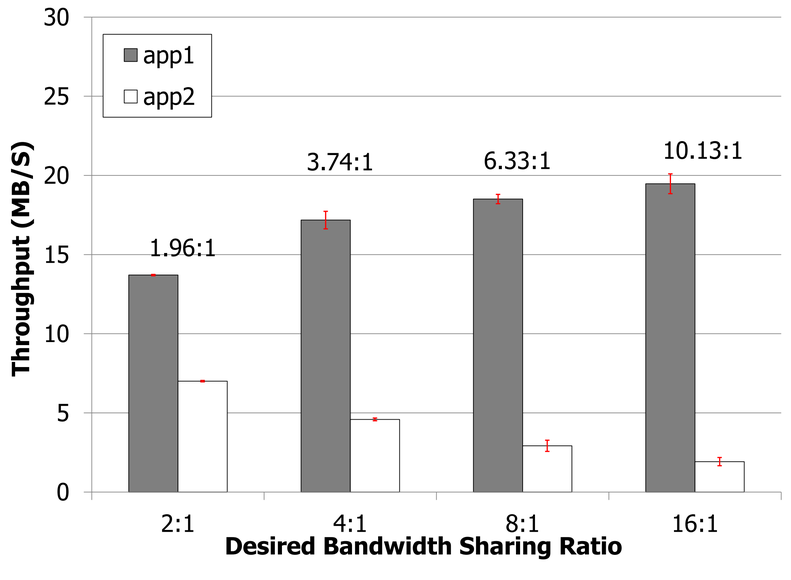
Prototype and Evaluation

- **PVFS2-based proxy prototype**
  - Intercept PVFS2 messages and virtualize a deployed physical PVFS2 system
  - Schedule parallel I/Os based on SFQ(D) — a proportional sharing algorithm
- **Experimental evaluation**
  - A virtual machine based testbed (up to 64 PVFS2 clients and 4 PVFS2 servers) hosted on eight physical cluster nodes
  - Two competing parallel applications (using the IOR benchmark) with large sequential writes

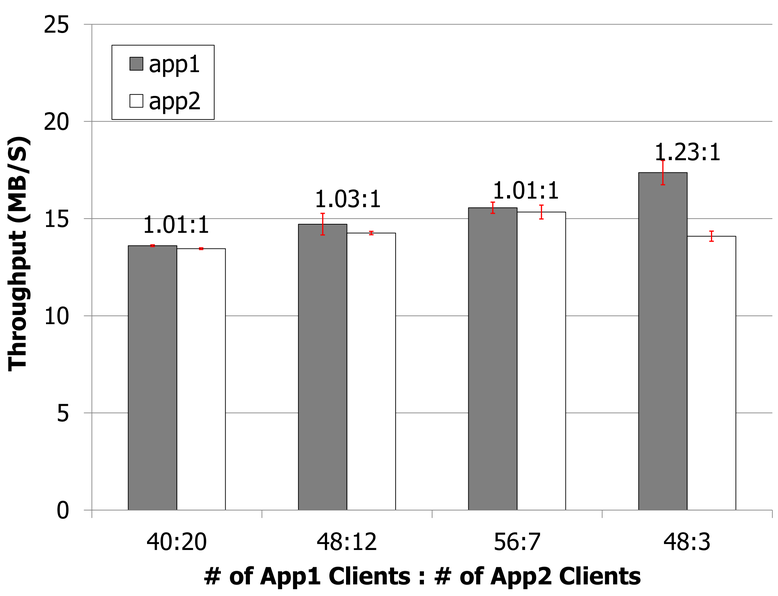
Proportional Sharing w/ Large I/Os



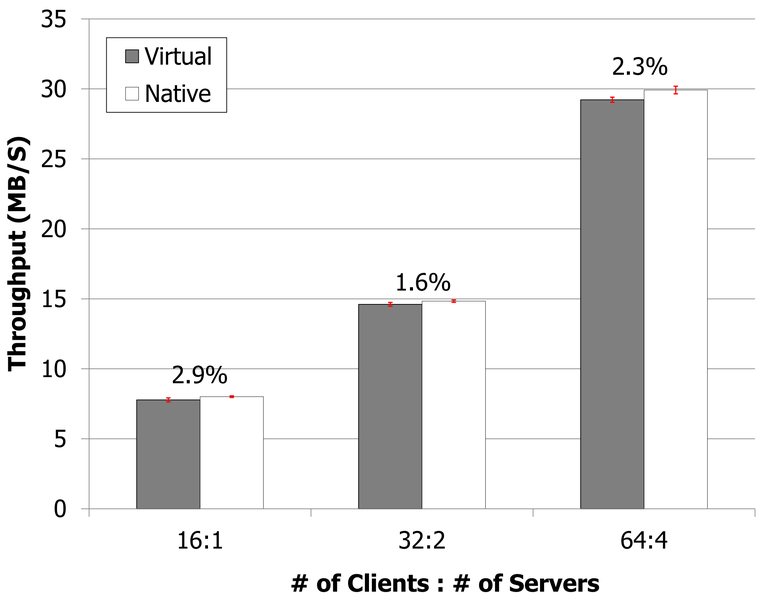
Proportional Sharing w/ Small I/Os



Prop. Sharing in Asymmetric Setup



Virtualization Overhead



Conclusion and Future Work

- **Conclusions**
  - PFS virtualization supports effective proportional bandwidth sharing of parallel storage systems
  - The performance and resource overhead of proxy-based PFS virtualization is small
- **Future Work**
  - Study the bandwidth management for applications with diverse I/O requirements
  - Study algorithms for global proportional sharing and deadline-driven I/O scheduling

FIU Virtualized Infrastructure, Systems, and Applications Research Laboratory (VISA)
http://visa.cis.fiu.edu

UF Advanced Computing and Information Systems Laboratory (ACIS)
http://www.acis.ufl.edu

Effort sponsored by National Science Foundation under grants CCF-0937973 and CCF-0938045. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.