QoS-driven Storage Management for High-end Computing Systems

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

Overview

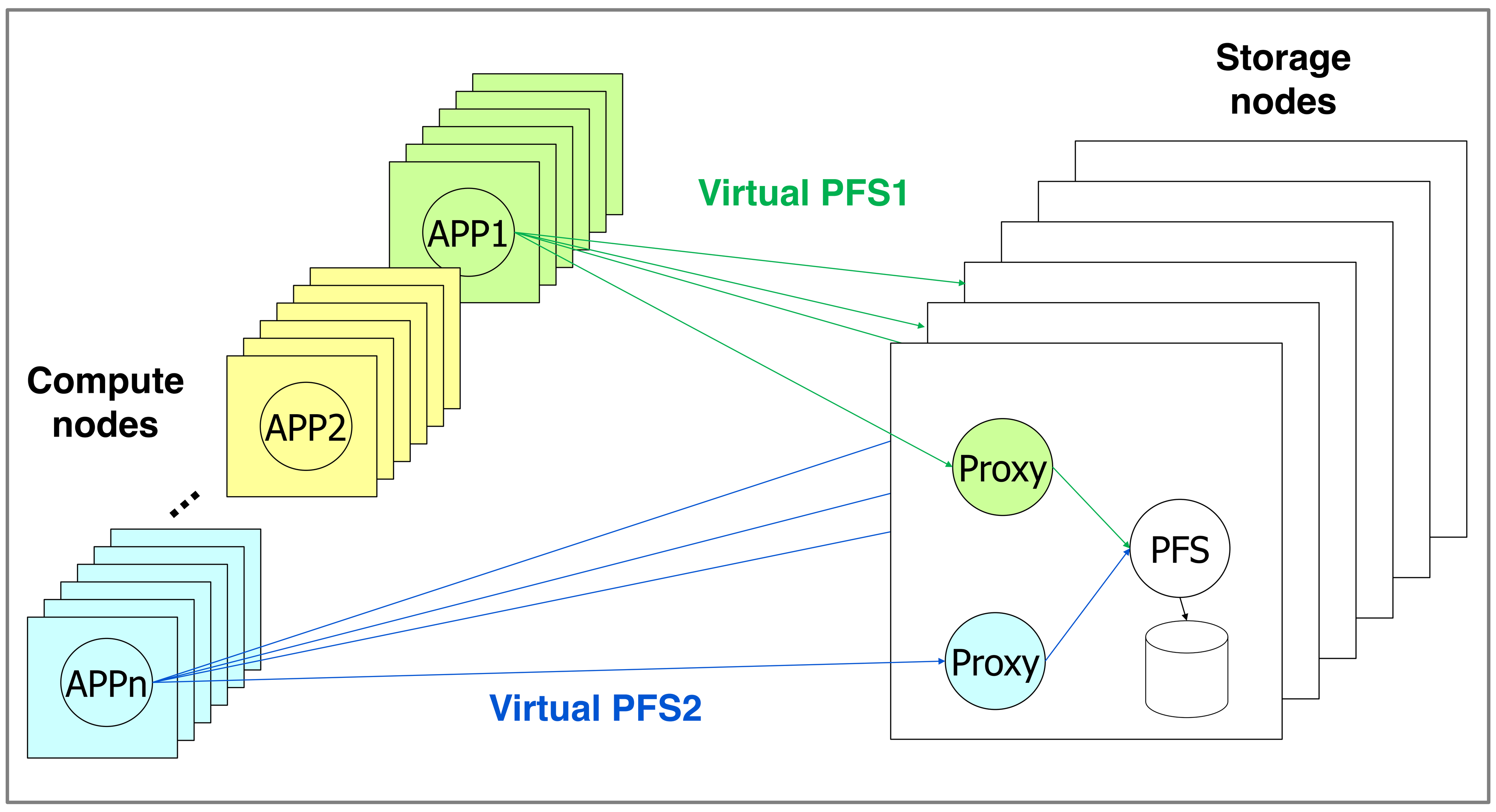
- **Goal:** Application-QoS driven storage resource management in high-end computing systems
- **Challenges:**
  - The lack of QoS differentiation in typical HEC parallel storage systems
  - The diversity in application I/O requirements
- **Solution:**
  - Parallel file system (PFS) virtualization based storage management
  - QoS-driven parallel I/O scheduling

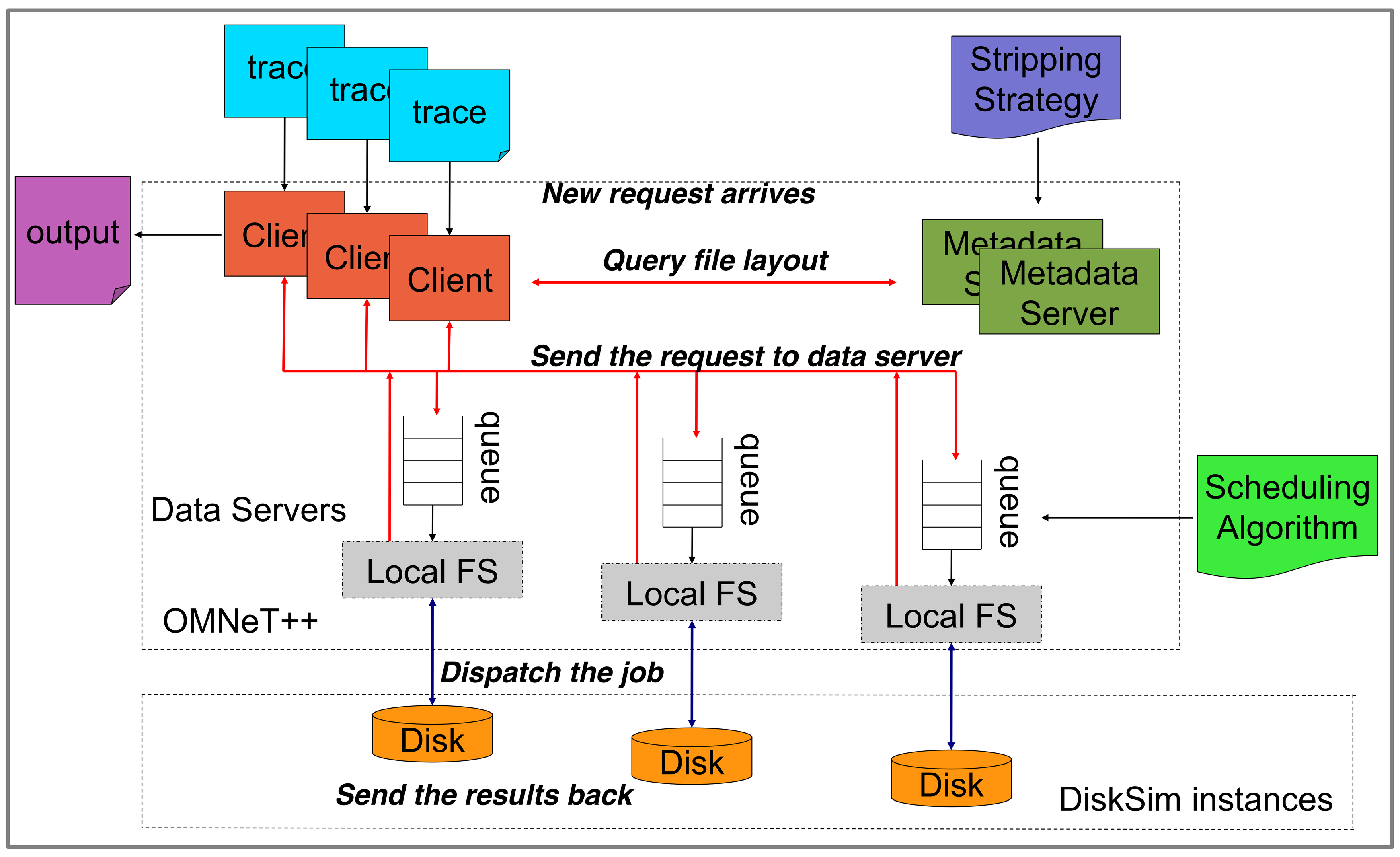
Parallel File System Virtualization

- **Per application virtual PFSes**
  - Dynamically created and destroyed based on application lifecycles
  - Application-specific I/O bandwidth allocation per virtual PFS
- **Proxy-based PFS virtualization**
  - Indirection of parallel I/Os between PFS clients and servers
  - Create per-application virtual PFSes and enforce I/O resource allocation

Simulation-based Scheduling Study

- **Parallel storage system simulator**
  - Flexibly study parallel I/O scheduling
  - Simulate enough system details but with an acceptable simulation time
- **Simulate PFS network**
  - Use discrete event simulation library (OMNeT++ 4.0)
- **Simulate PFS disks**
  - Use disk system simulator (DiskSim 4.0)

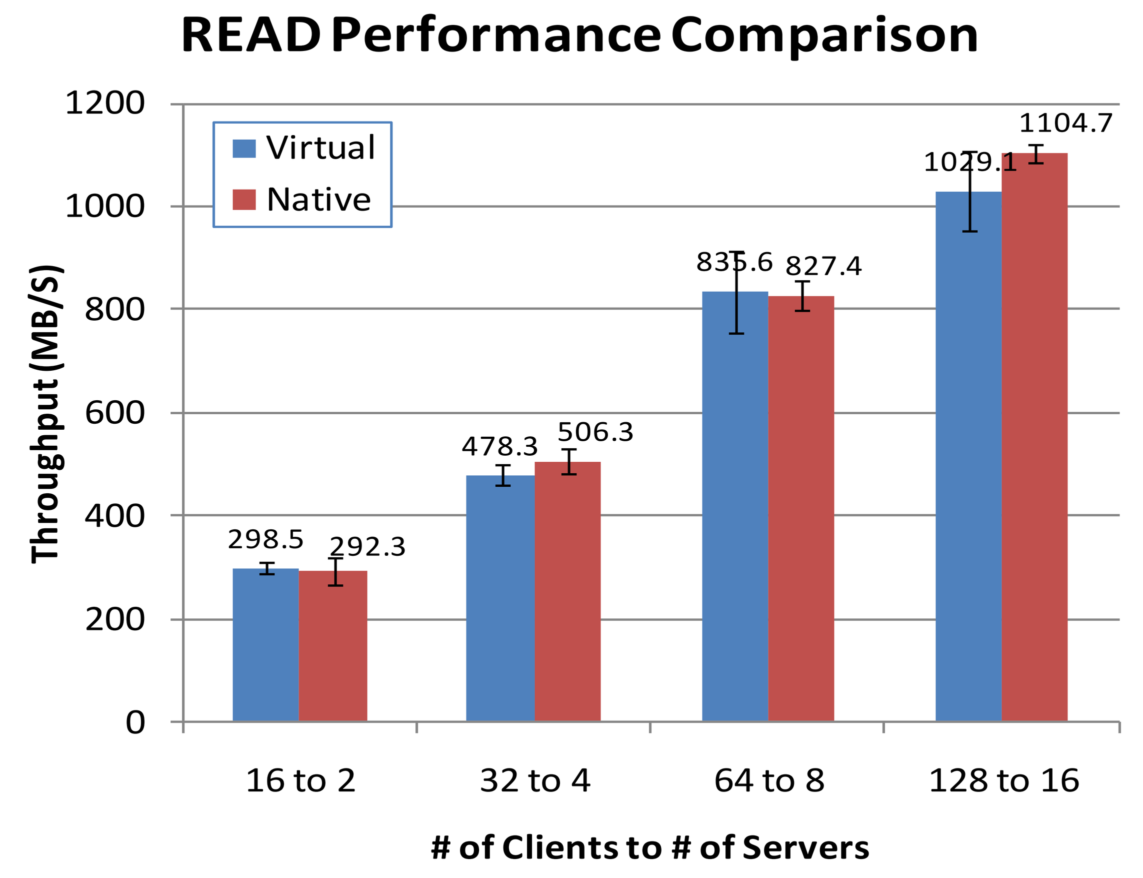


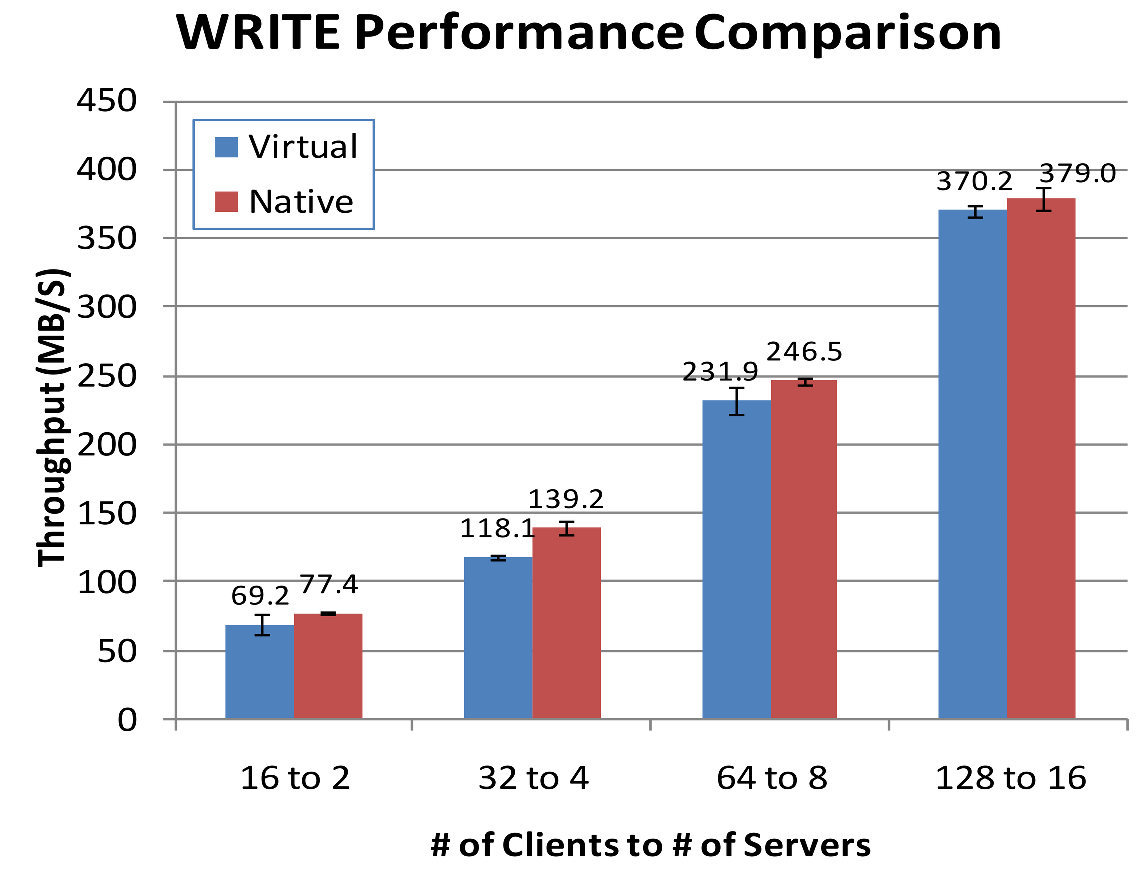


Implementation and Evaluation

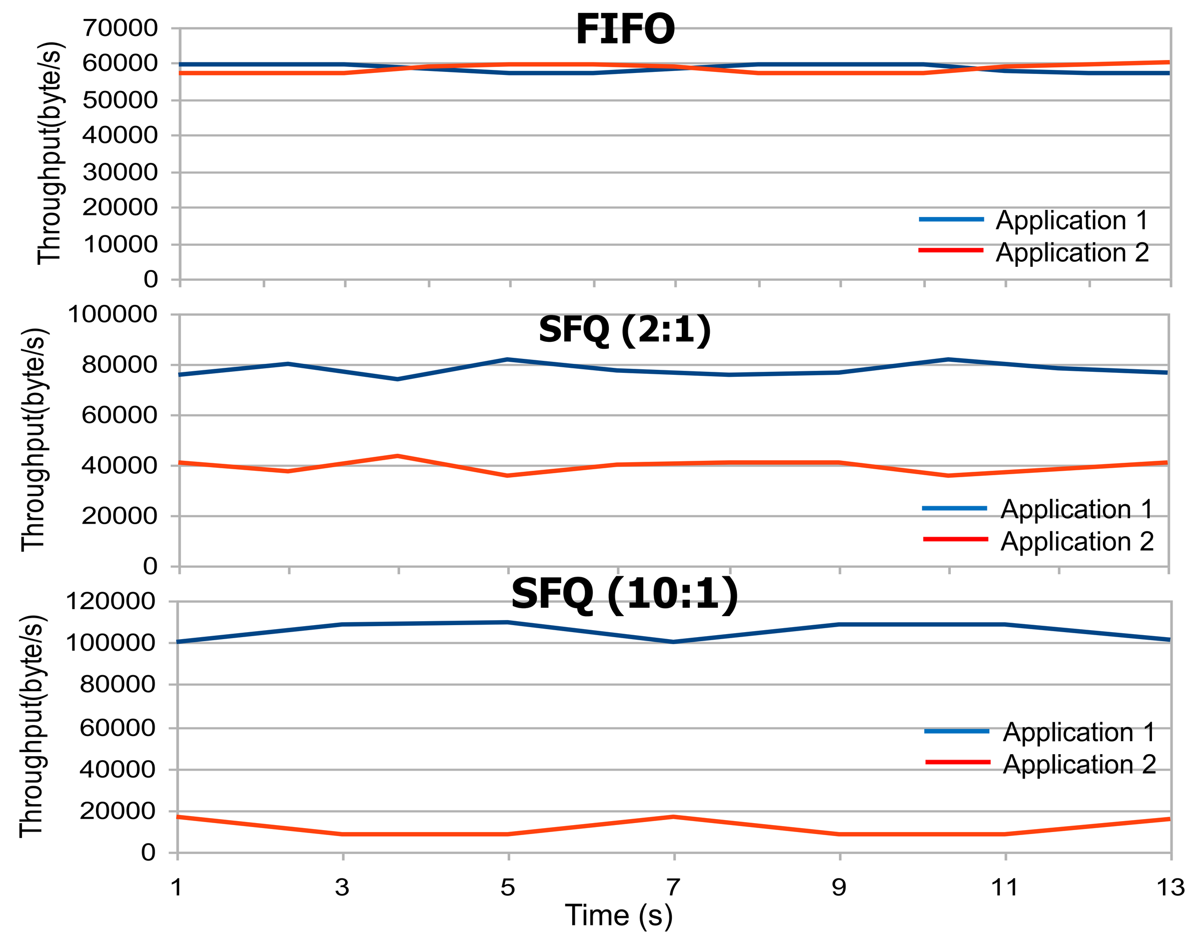
- **PVFS2 (Parallel Virtual File System) proxy**
  - Intercept PVFS2 messages and virtualize a deployed physical PVFS2 system
- **Evaluation**
  - A virtual machine based testbed (Up to 128 PVFS clients and 16 PVFS servers)
  - Benchmark: IOR version 2
- **Simulation**
  - Two parallel applications (16 clients each)
  - Four data servers
  - Driven by traces generated from IOR

Virtualization Overhead



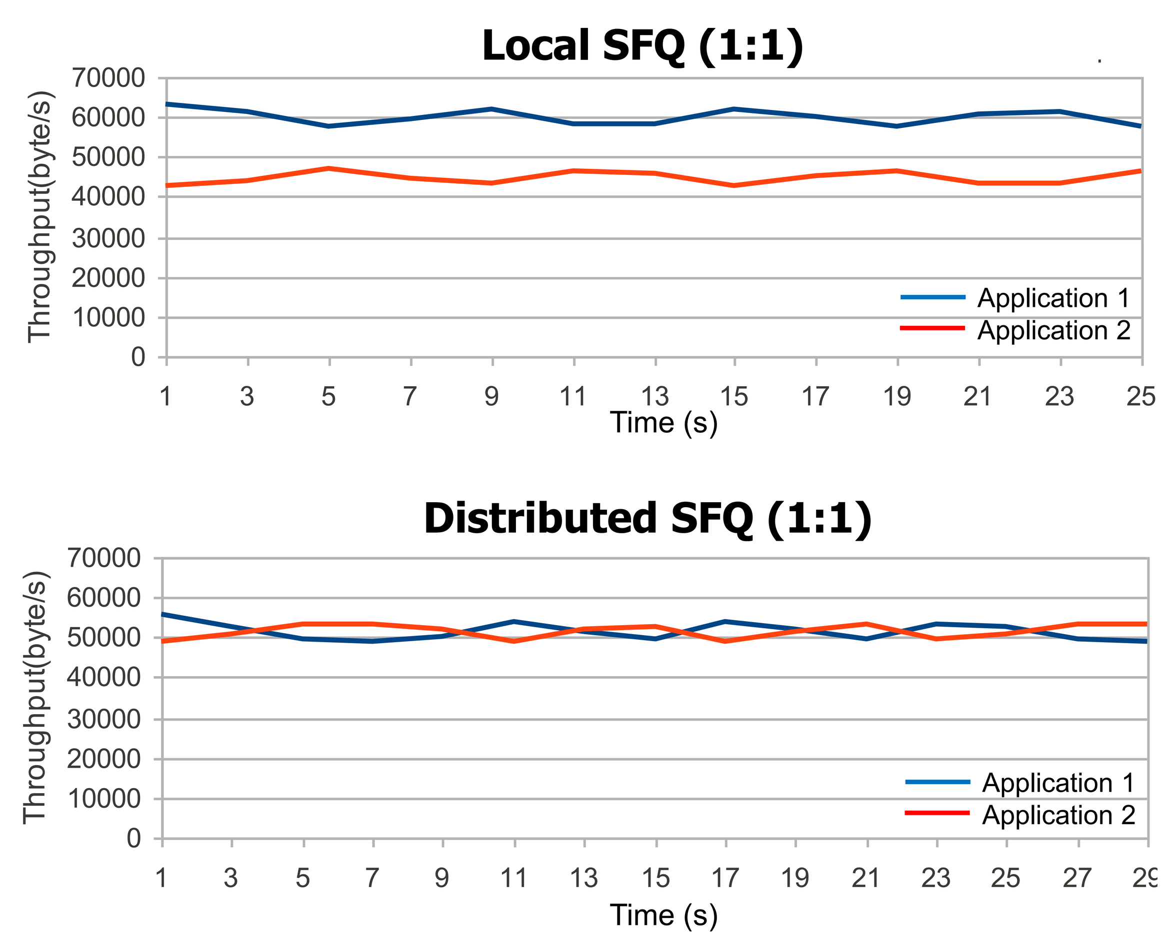


Simulation: FIFO vs. SFQ



Application 1 & 2 share all the four data servers

Simulation: Local SFQ vs. DSFQ



App 1: four data servers; App 2: three data servers

Conclusion and Ongoing Work

- **Proxy-based PFS virtualization is feasible**
  - Its throughput overhead and resource usage overhead are not significant
  - TODO: implement optimized I/O schedulers upon proxy
- **Simulation-based PFS scheduling study is valuable**
  - Its results can guide the design of real I/O schedulers
  - TODO: improve the scale and realism of simulation

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.