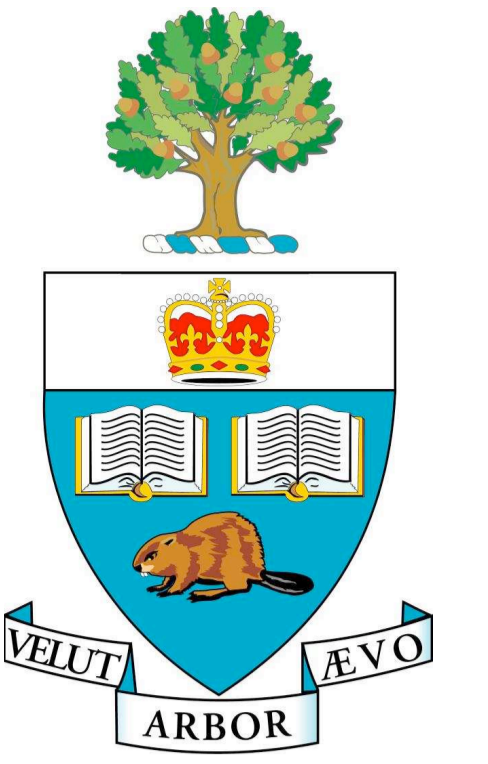


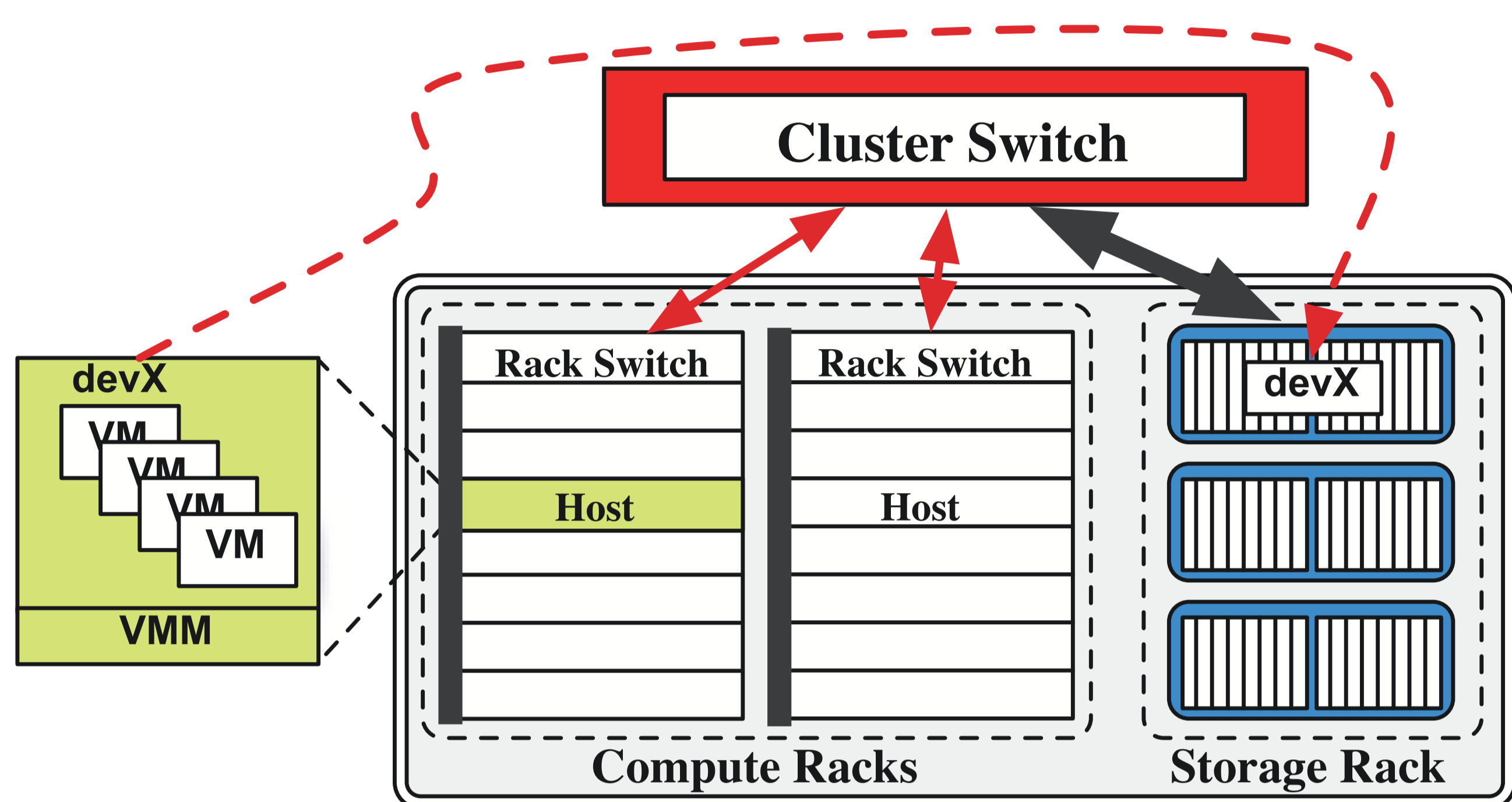
SLIM: End-to-End I/O Decongestion for Network Storage

Madalin Mihailescu[†], Gokul Soundararajan, Cristiana Amza

Department of Computer Science[†]
Department of Electrical and Computer Engineering
University of Toronto



Virtualized Data Centers



- Types of Network Storage
 - *Dedicated*: High cost/high performance - Fibre Channel
 - *Shared*: Low cost/low performance - iSCSI/Ethernet
- **Goal**: Increase performance for low-cost network storage

Overview

- **Problem**: Increasing I/O needs in data centers
 - Data intensive applications: Map-Reduce, Hadoop
 - Low-cost consolidated network storage
 - *Result*: Rack-external links **oversubscribed**
- **Observation**: Abundance of resources *within racks*
 - Underutilized disks on each server - SSDs
 - Underutilized computation power - multi-core CPUs
 - Underutilized network bandwidth - 10Gbps Ethernet
- **Idea**: Use rack-level resources to reduce network storage traffic

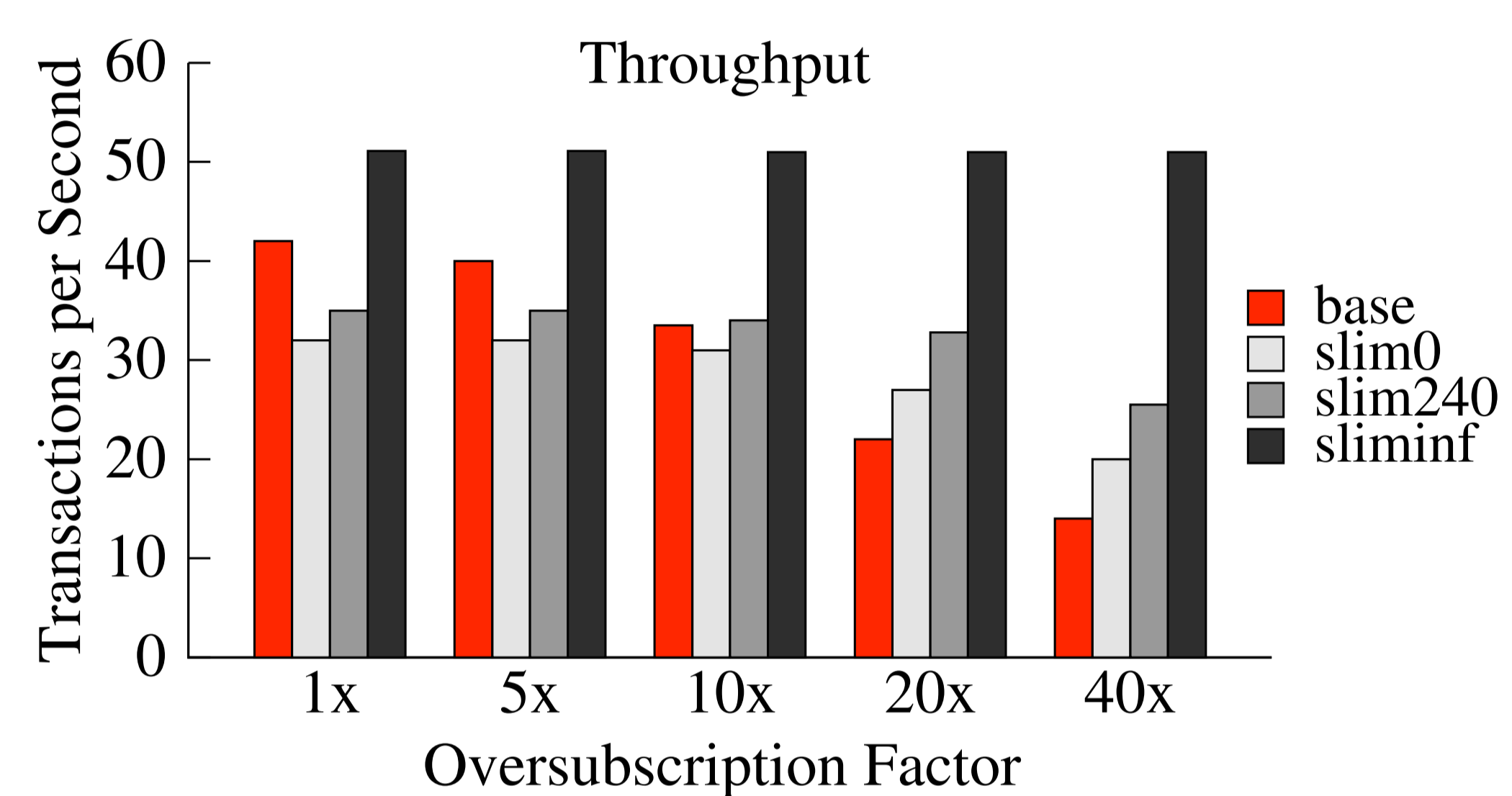
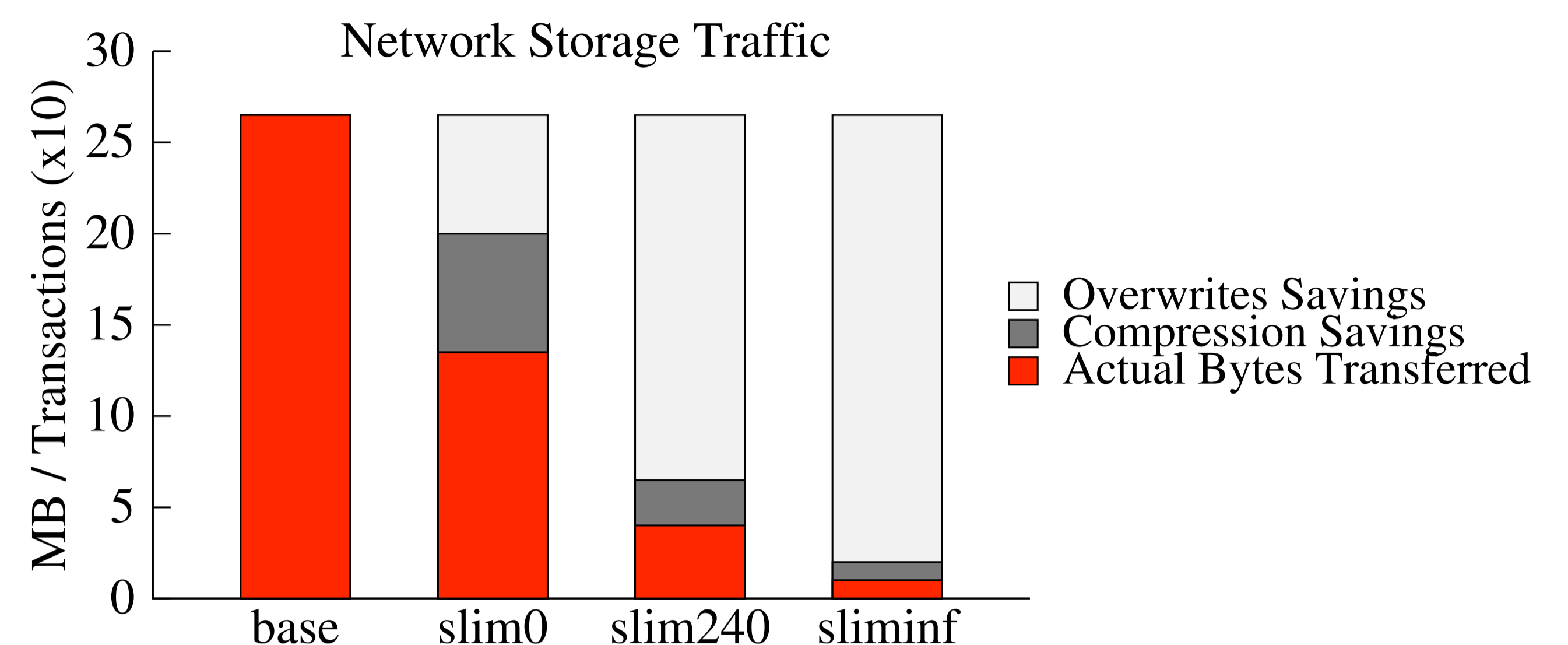
Approach

- **Build rack-level persistent write-back cache**
 - Pool server disks to build persistent *rack cache*
 - Replicate within rack for high availability
 - Transparently offload I/O writes to rack cache
 - Periodically migrate data from rack cache to network storage
- **Enable I/O optimizations**
 - Leverage I/O overwrite patterns
 - Leverage CPU for compression outside critical path

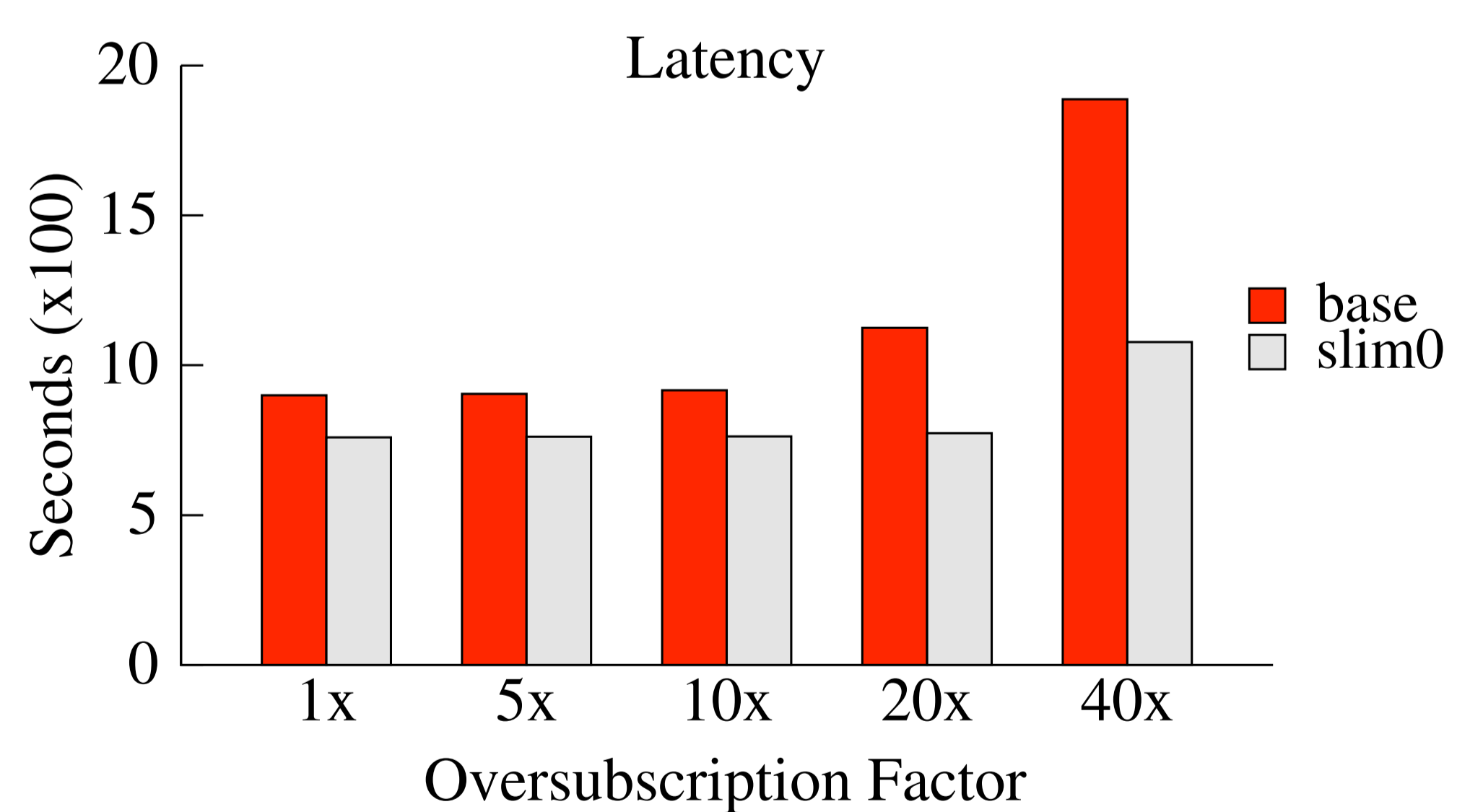
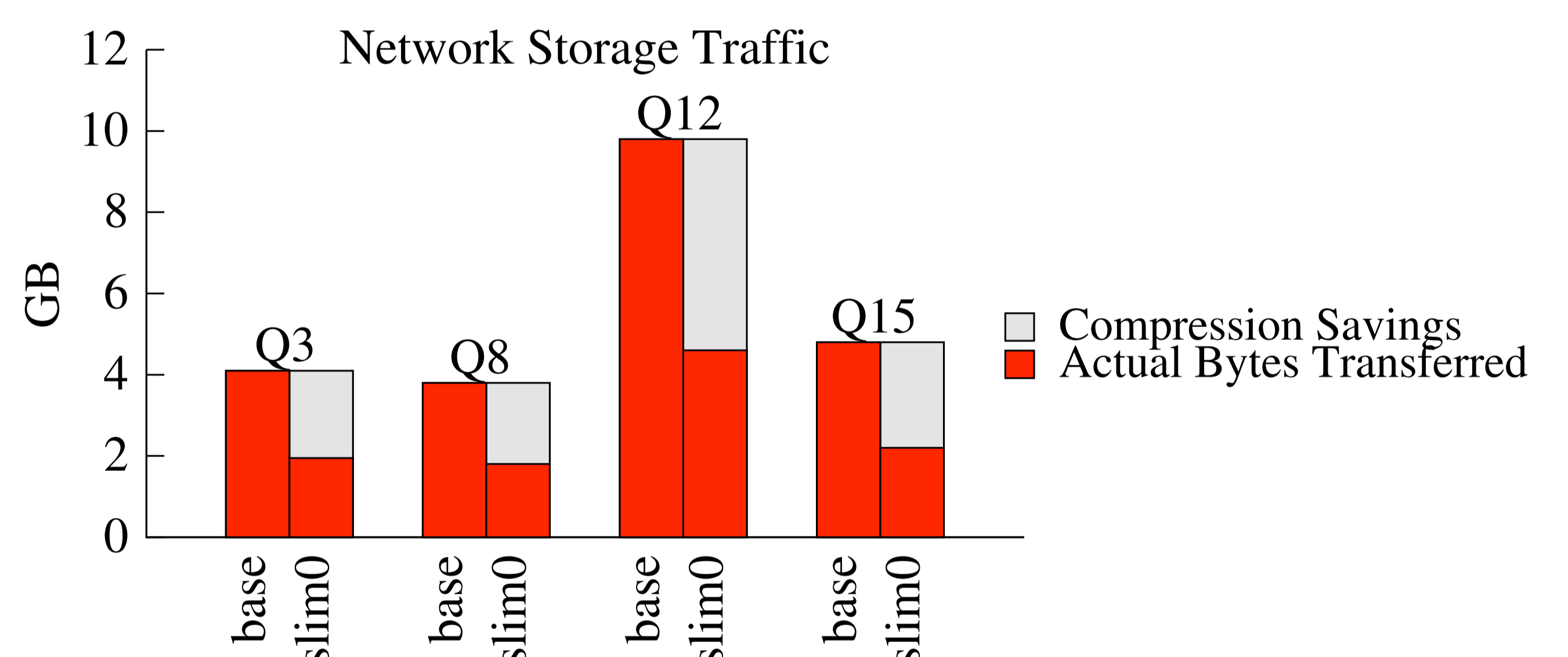
Prototype

- **Transparent to VM client and network storage server**
 - Built using Network Block Device (NBD); zlib for compression
 - Store data compressed in network storage
- **Optimized for performance**
 - Log-structured storage for rack cache
 - Decompression on the fly for I/O reads

Write-Intensive Workload – TPC-C



Read-Intensive Workload – TPC-H



Ongoing Work

- **Improving current prototype**
- **Evaluating for different scenarios**
 - Measuring scalability in large clusters
 - Studying adaptiveness to client I/O patterns