Simulation Study of iSCSI-Based Storage

Yingping Lu, Farrukh Norman & David Du

(lu, norman, du)@cs.umn.edu

Digital Technology Center Intelligent Storage Consortium

University of Minnesota

NASA/IEEE MSST 2004

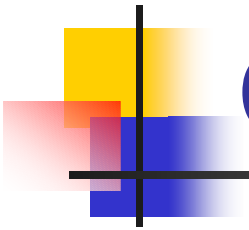
12th NASA Goddard/21st IEEE Conference on Mass Storage Systems & Technologies

The Inn and Conference Center

University of Maryland University College

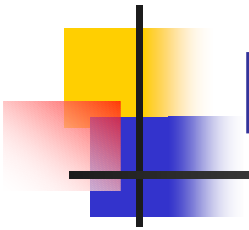
Adelphi MD USA

April 13-16, 2004



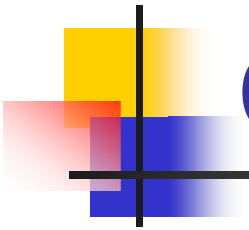
Outline

- Motivation
- Simulation Model
- Implementation
- Performance Evaluation
- Conclusion



Motivation

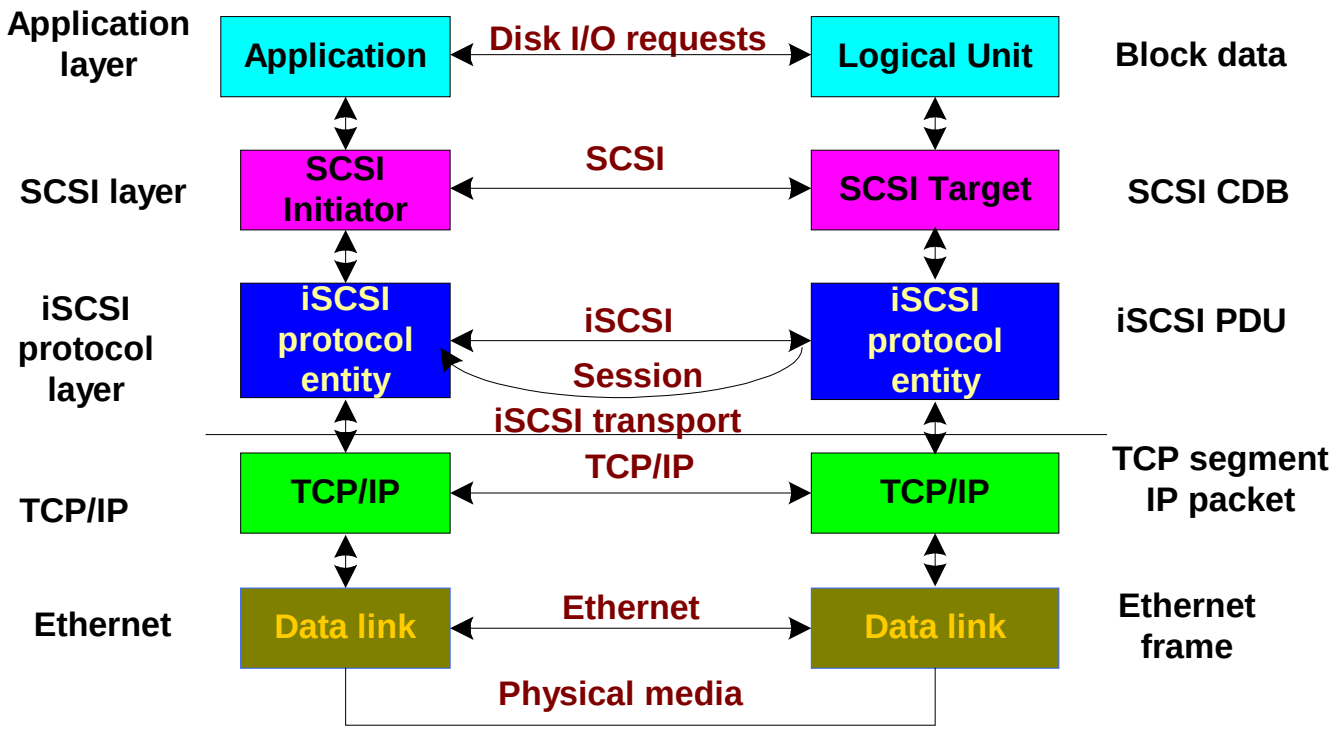
- iSCSI creates a new storage paradigm
  - Greatly extend the storage distance
  - Exploit the ubiquity of the Internet
- The underlying TCP/IP protocol has a lot of uncertainty
  - TCP/IP is an open protocol
  - The network infrastructure is heterogeneous
- A performance tool to assist:
  - The evaluation of design alternatives and tradeoffs
  - The study of performance characteristics and developing of new applications, etc.



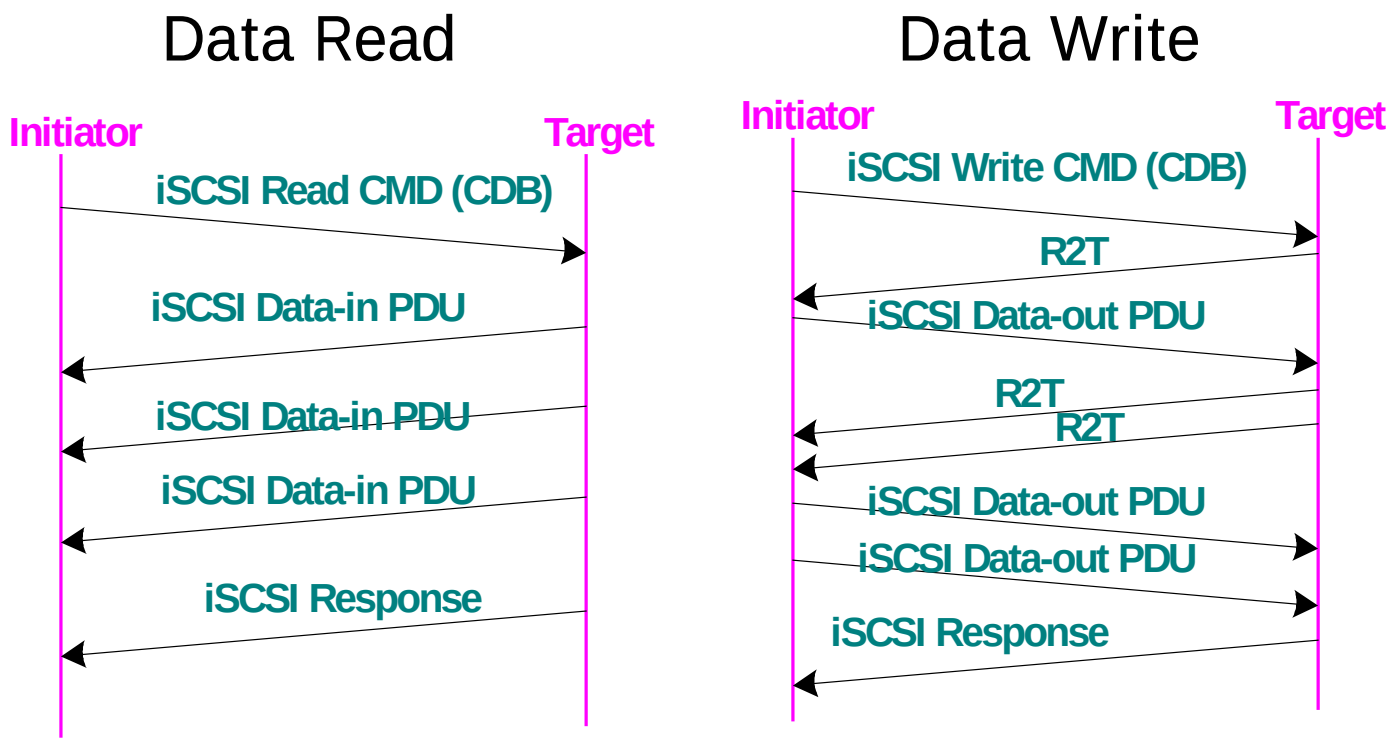
Objectives

- Exploit the existing NS-2 simulation tools developed for TCP/IP network
  - A variety of built-in protocols, flow control mechanisms, and flexible configurations
- Create a generic simulation tool consisting network and storage components for iSCSI
  - Modular and well-defined interface between components
  - Easy configuration of test setting
- Study the impact of network setting to the storage access performance
  - The PDU length
  - The network bandwidth and delay

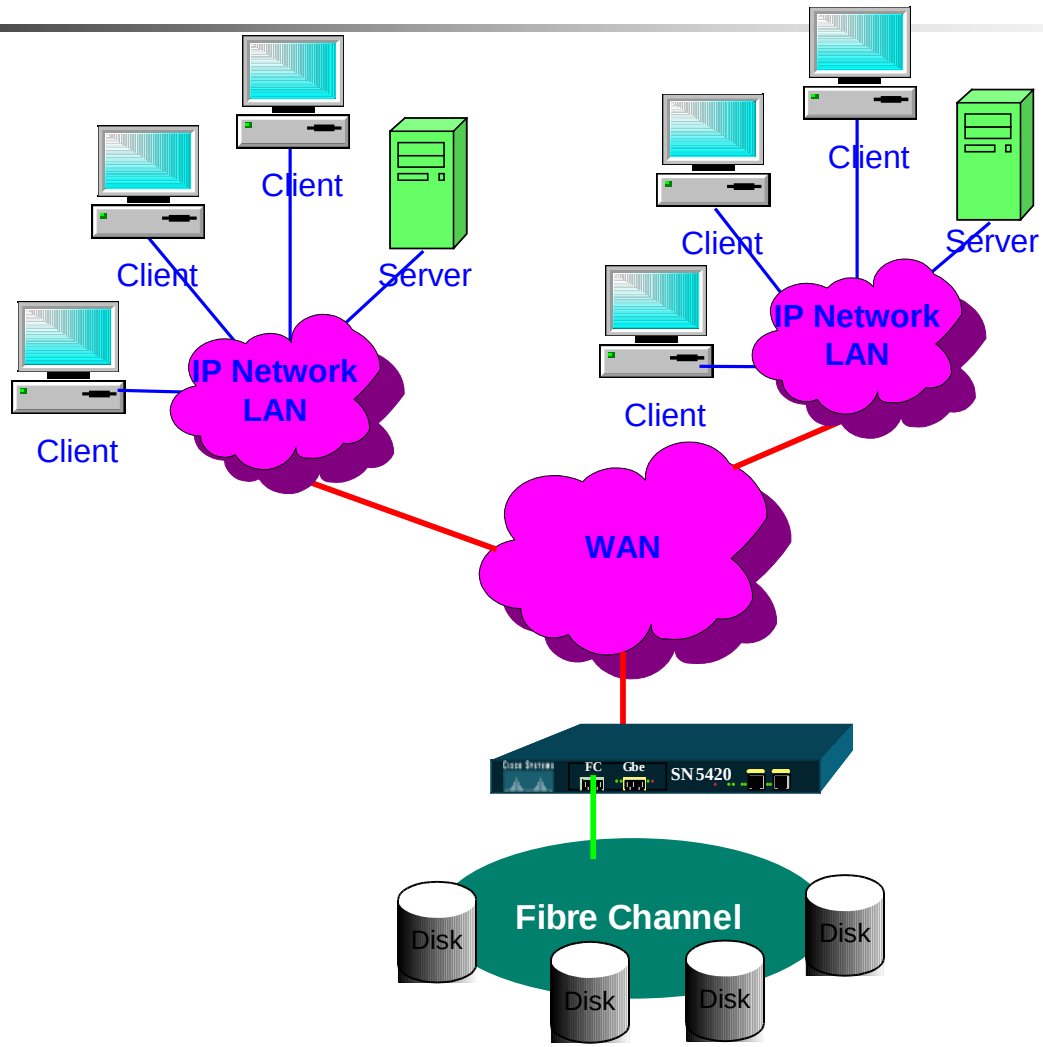
iSCSI Protocol Structure

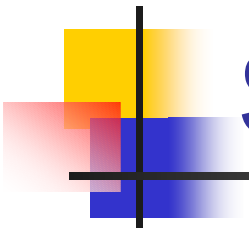


Data Access Operation



System Model

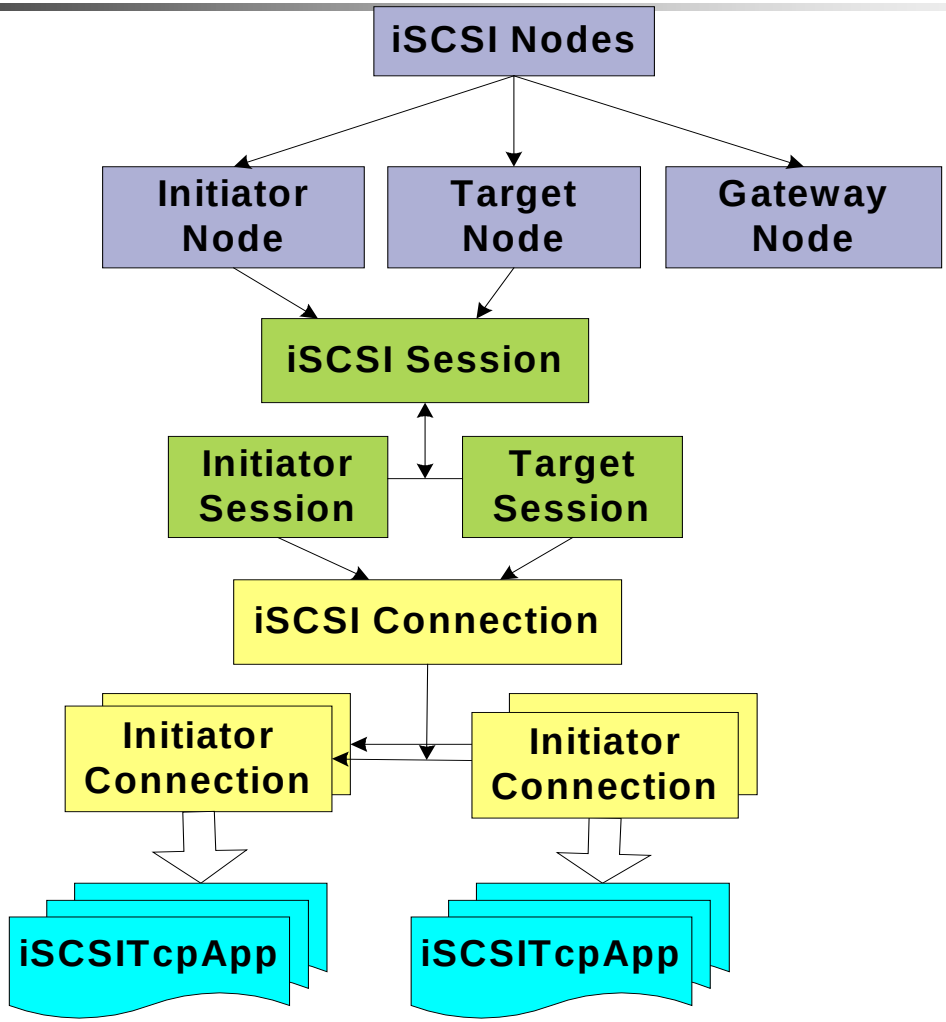




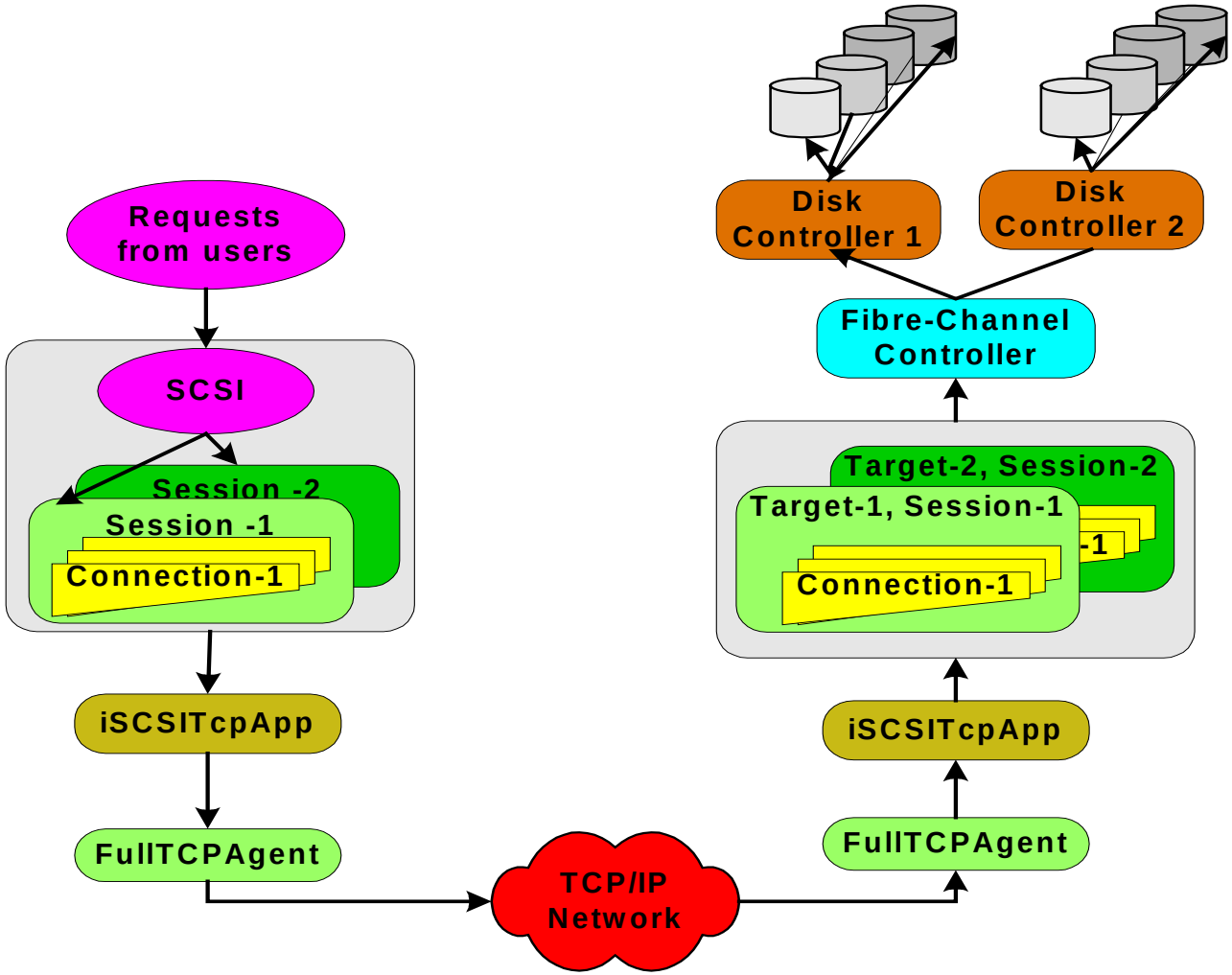
Simulation Components

- Initiator node
- Storage Gateway
- Target
- TCP Agent
- FC-Channel Link
- Disk

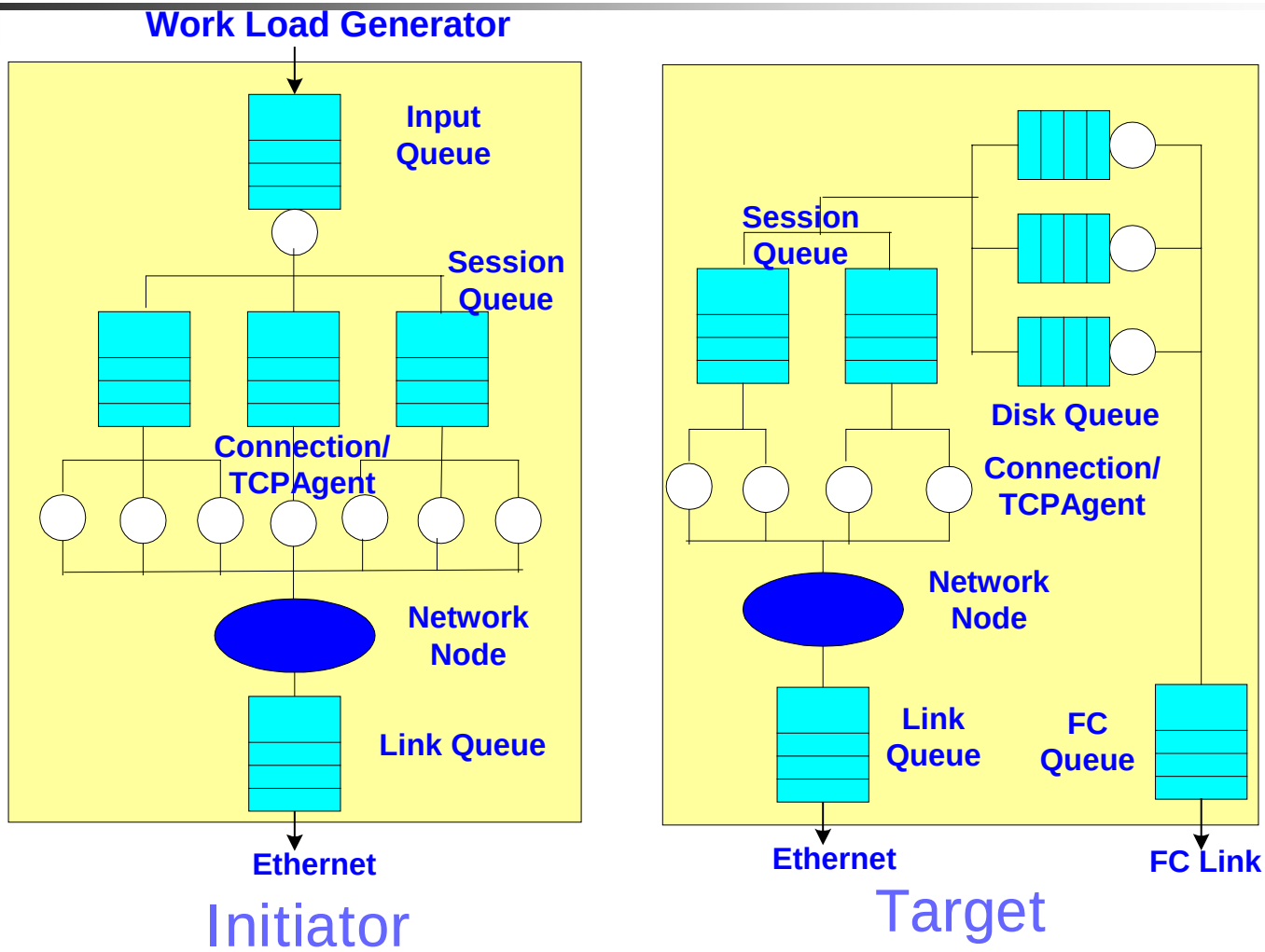
Implementation: iSCSI Node



System Architecture

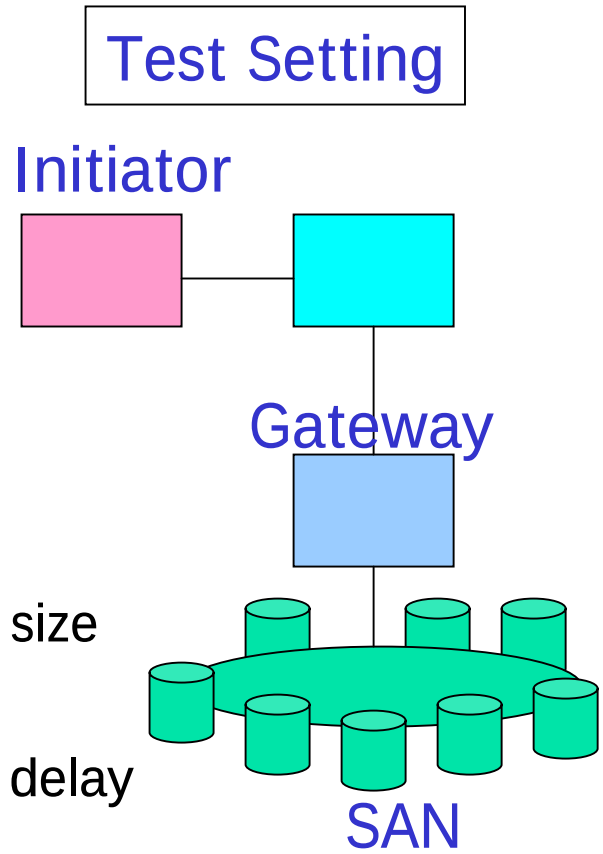


Queuing Model

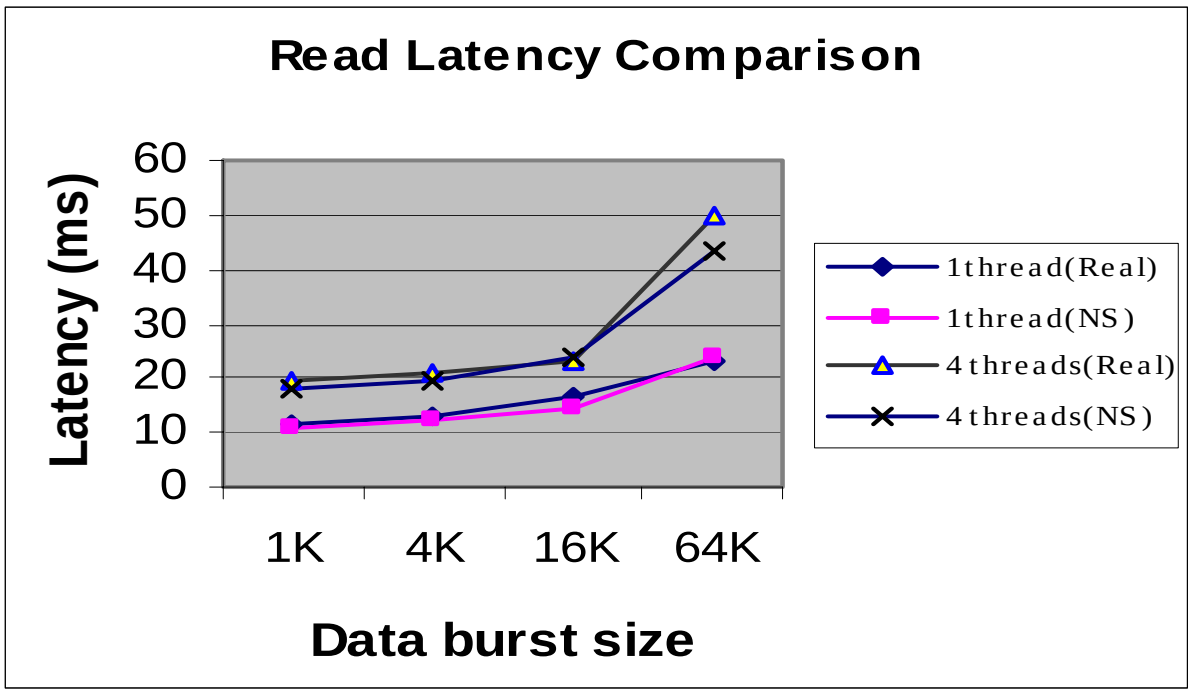


Simulation Setting

- Disk Model:
  - Seagate ST39102FC Cheetah 9LP
  - Rotation speed: 10025 RPM
- Fibre Channel:
  - 1Gb/s link
- Gateway node
  - 1 target
  - 8 disks in the target
  - Adjustable window size, segment size
- TCP/IP network
  - 1 link with adjustable bandwidth, delay

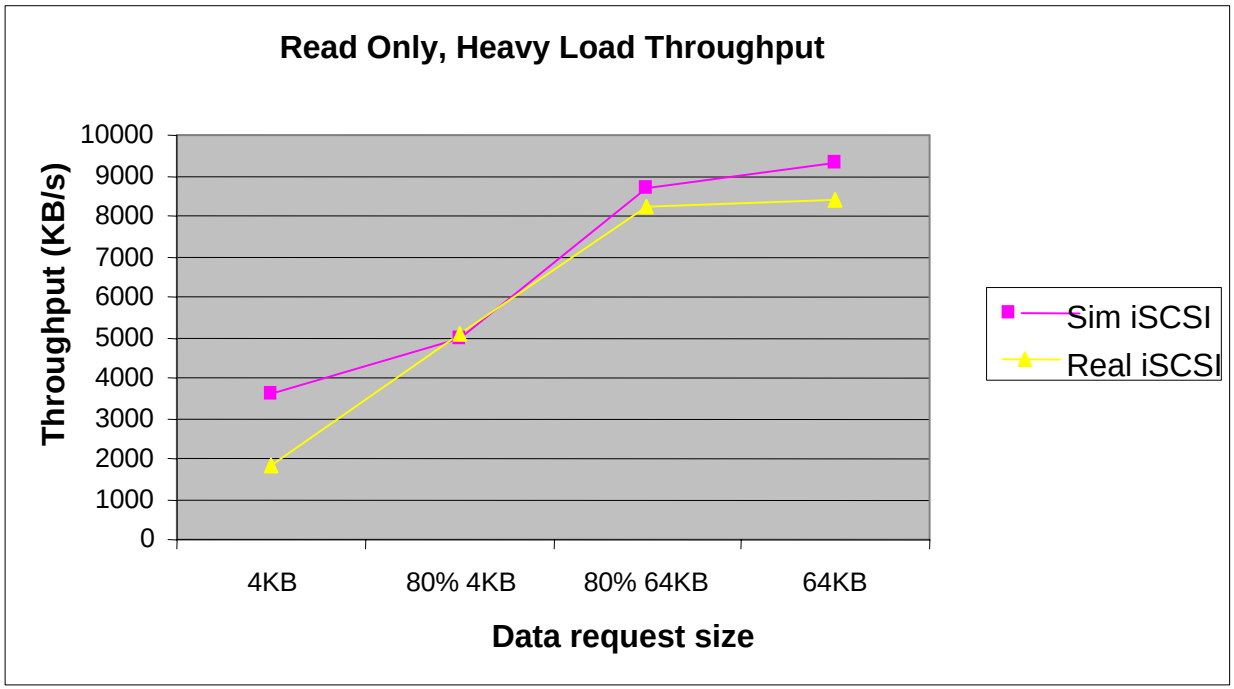


Latency Comparison



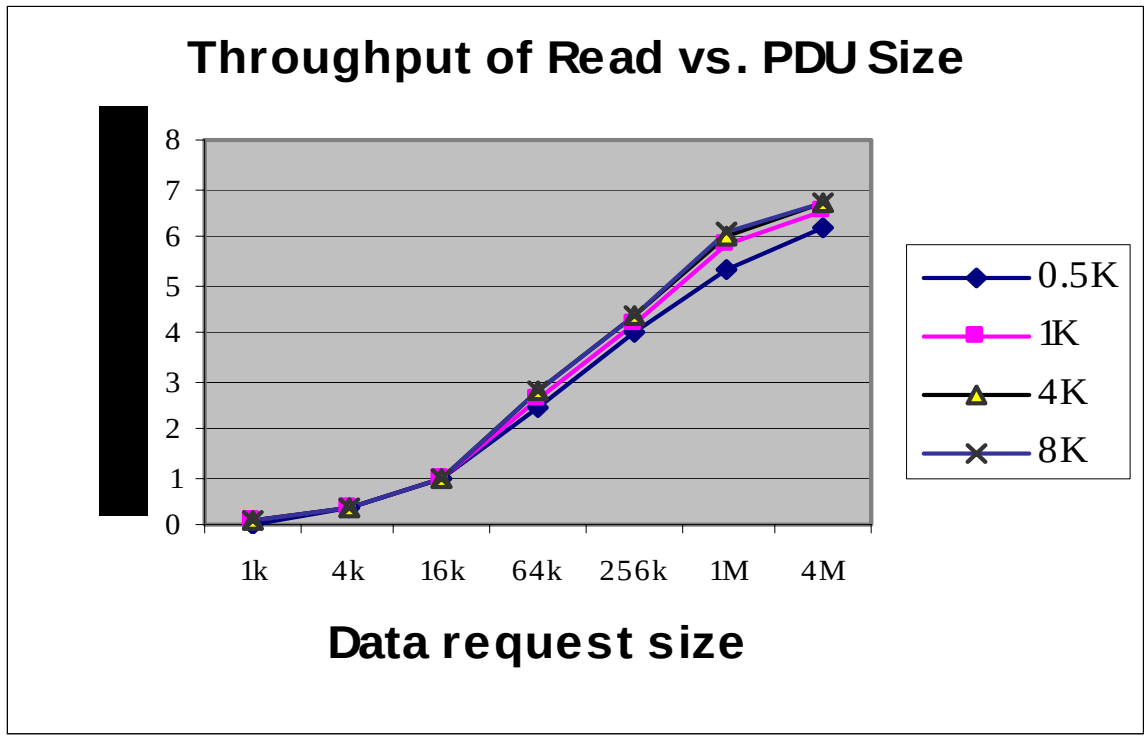
Read
PDU=8KB
Delay=1.5ms
Win=80
Link BW=100Mb/s

Throughput Comparison



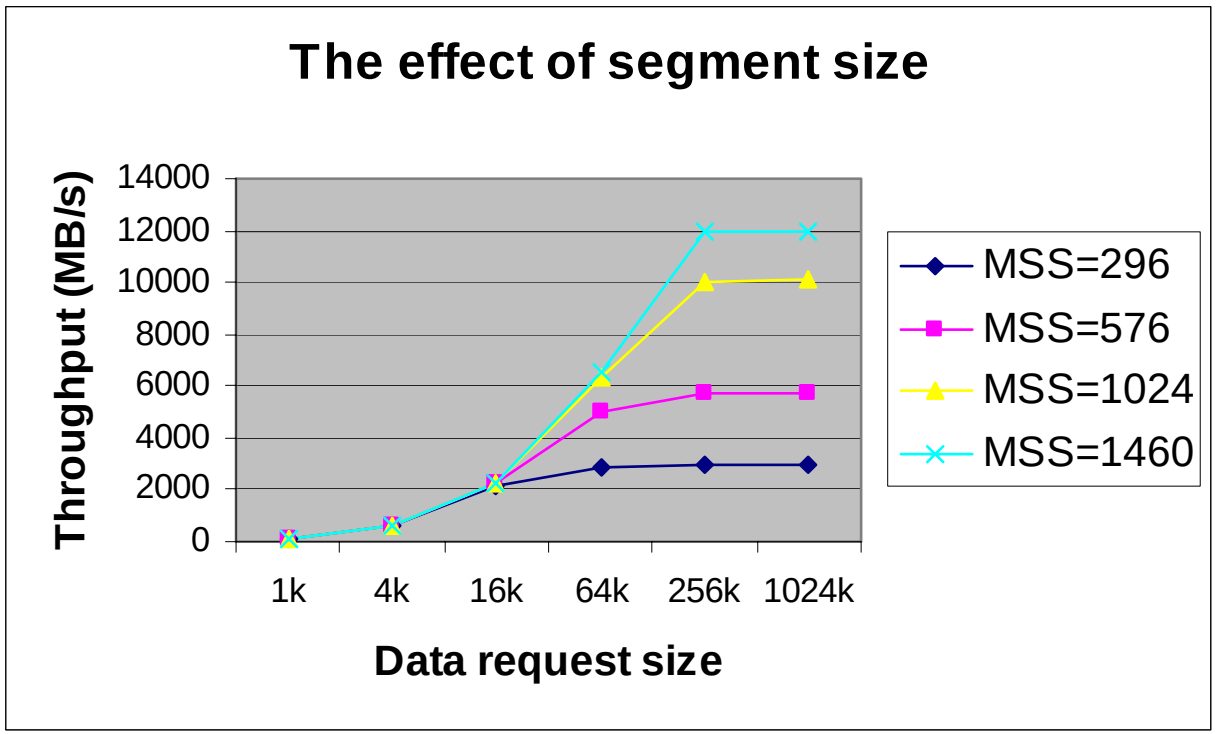
Read
PDU=8KB
Delay=1.5ms
Win=80
Link BW=100Mb/s

Performance Evaluation



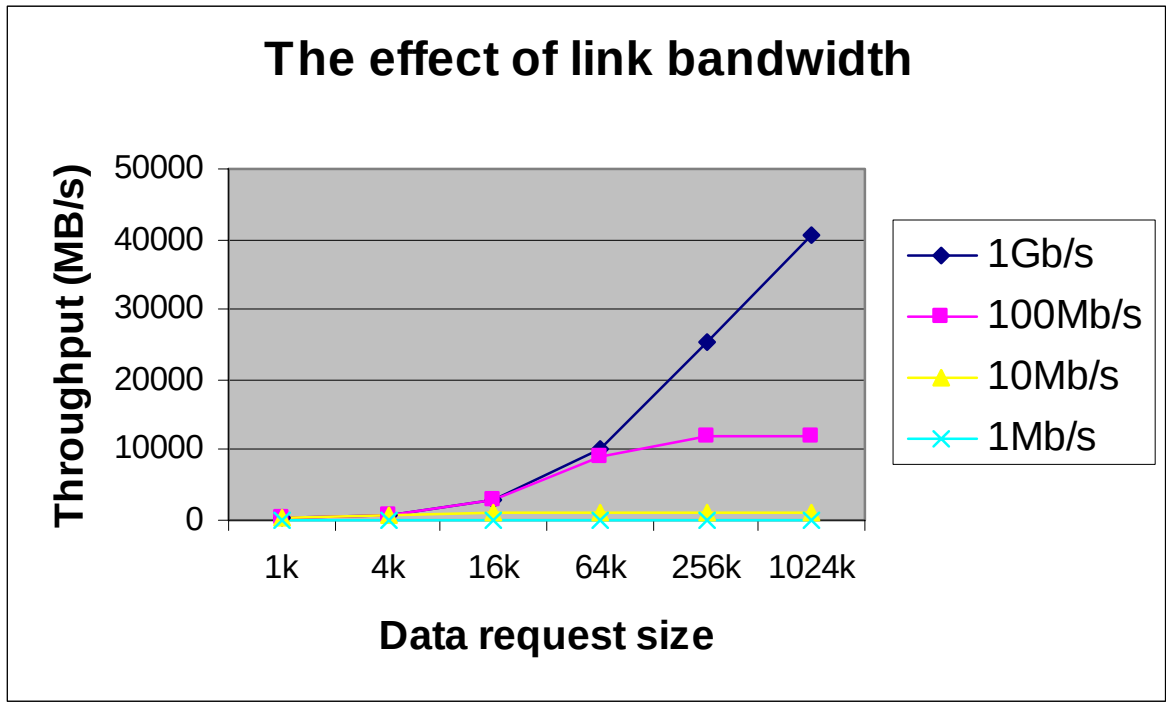
Read
PDU=8KB
#Thread=1
Delay=1.5ms
Win=80
Link BW=100Mb/s

The Effect of Segment Size



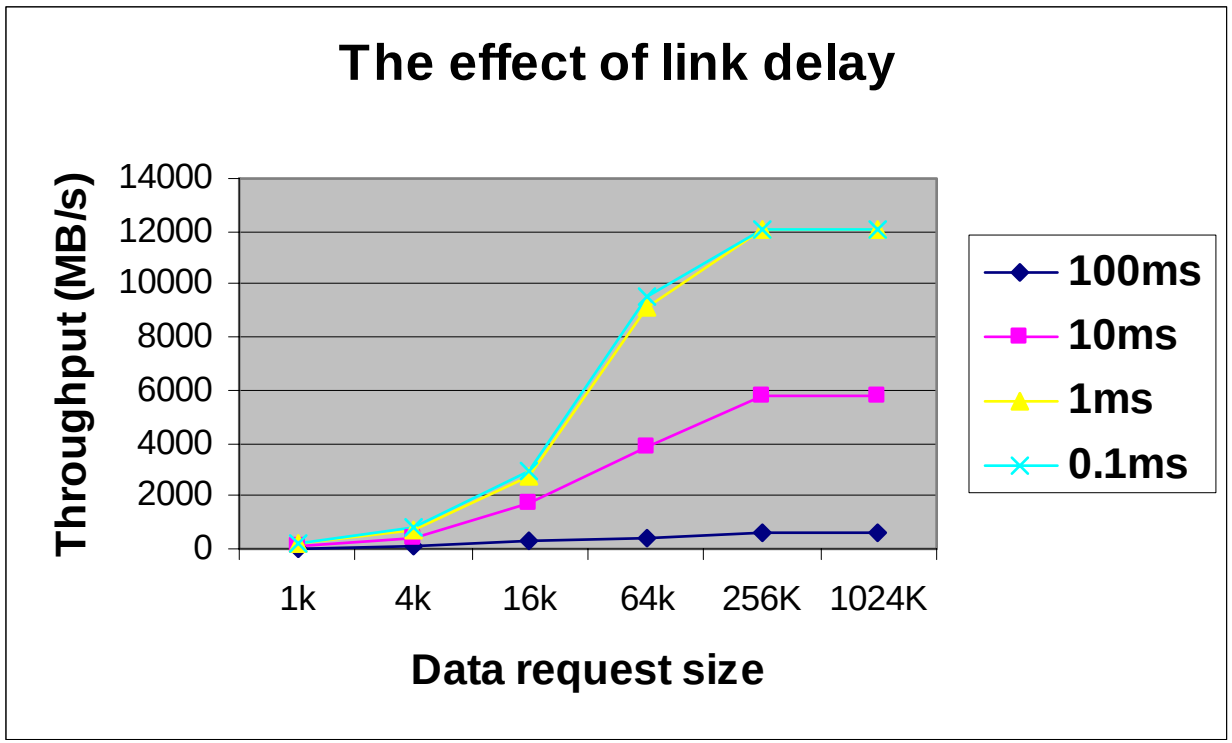
Write
PDU=8KB
#Thread=4
Delay=5ms
Win = 100
Link BW=100Mb/s

The Effect of Link Bandwidth



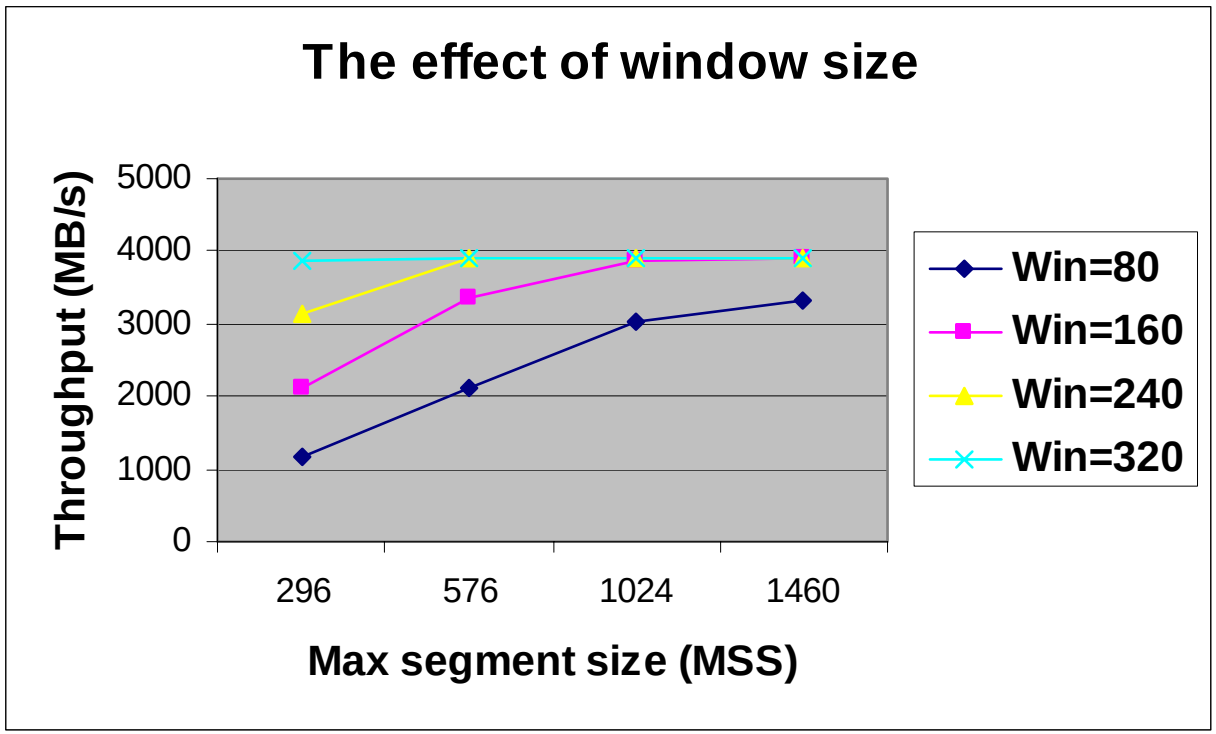
Write
PDU=8KB
Win = 80
#Thread=4
Delay=1ms
MSS=1460

The Effect of Link Delay

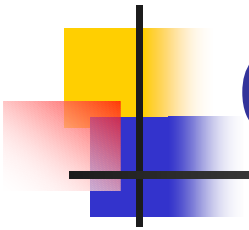


Write
PDU=8KB
MSS=1460
#Thread=4
Win = 80
Link BW=100Mb/s

The Effect of Window Size

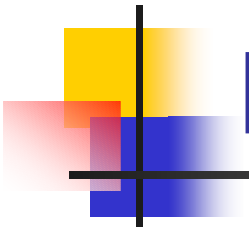


Write
PDU=8KB
MSS=1460
#Thread=4
Data size = 64KB
Link BW=100Mb/s



Conclusion

- Integrate the TCP/IP network and storage simulation
- Study the impact of network characteristics to the performance of iSCSI storage system
  - The impact of PDU size
  - The impact of link delay
  - The impact of network link bandwidth
  - The impact of window size



Future Work

- Study the effect of TCP flow control mechanism, error control
- Apply scheduling algorithm and caching scheme in disk
- Implement storage brick and RAID function in the target