Data Storage Institute

An iSCSI Design and implementation

Hui Xiong

Tel:+0065-68748100

e_mail:Xiong_Hui@dsi.a-star.edu.sg

NASA/IEEE MSST 2004
12th NASA Goddard/21st IEEE Conference on
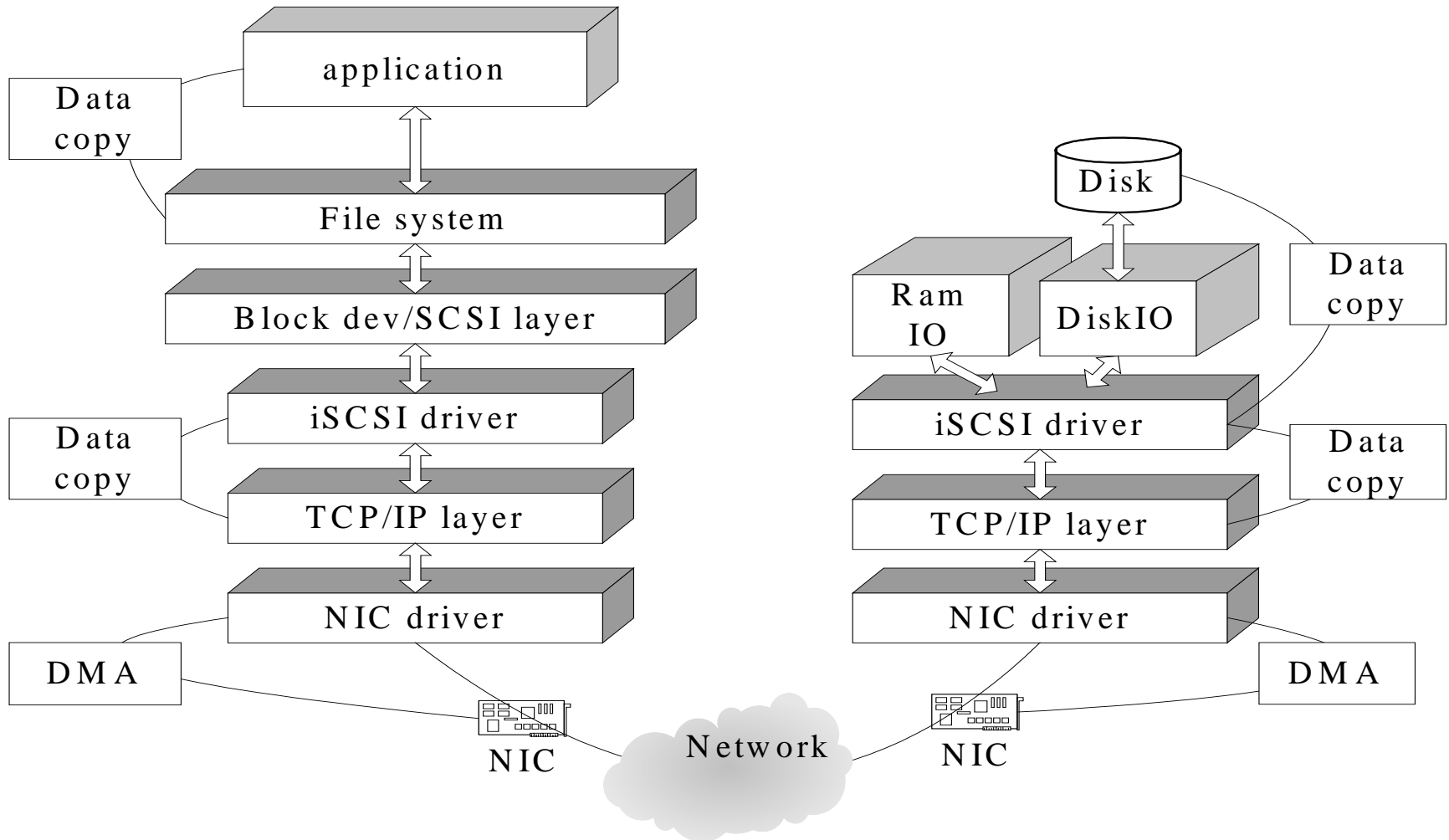Mass Storage Systems & Technologies
The Inn and Conference Center
University of Maryland University College
Adelphi MD USA
April 13-16, 2004

General Architecture of iSCSI

application

Data copy

File system

Block dev/SCSI layer

iSCSI driver

Data copy

TCP/IP layer

NIC driver

DMA

NIC

Network

Disk

Ram IO

DiskIO

Data copy

iSCSI driver

Data copy

TCP/IP layer

NIC driver

DMA

NIC

Data Storage Institute

A*STAR

Problem of iSCSI

Poor performance

- Client / Server need to handle TCP/IP & iSCSI stack

Some solutions

- TCP/IP offloading
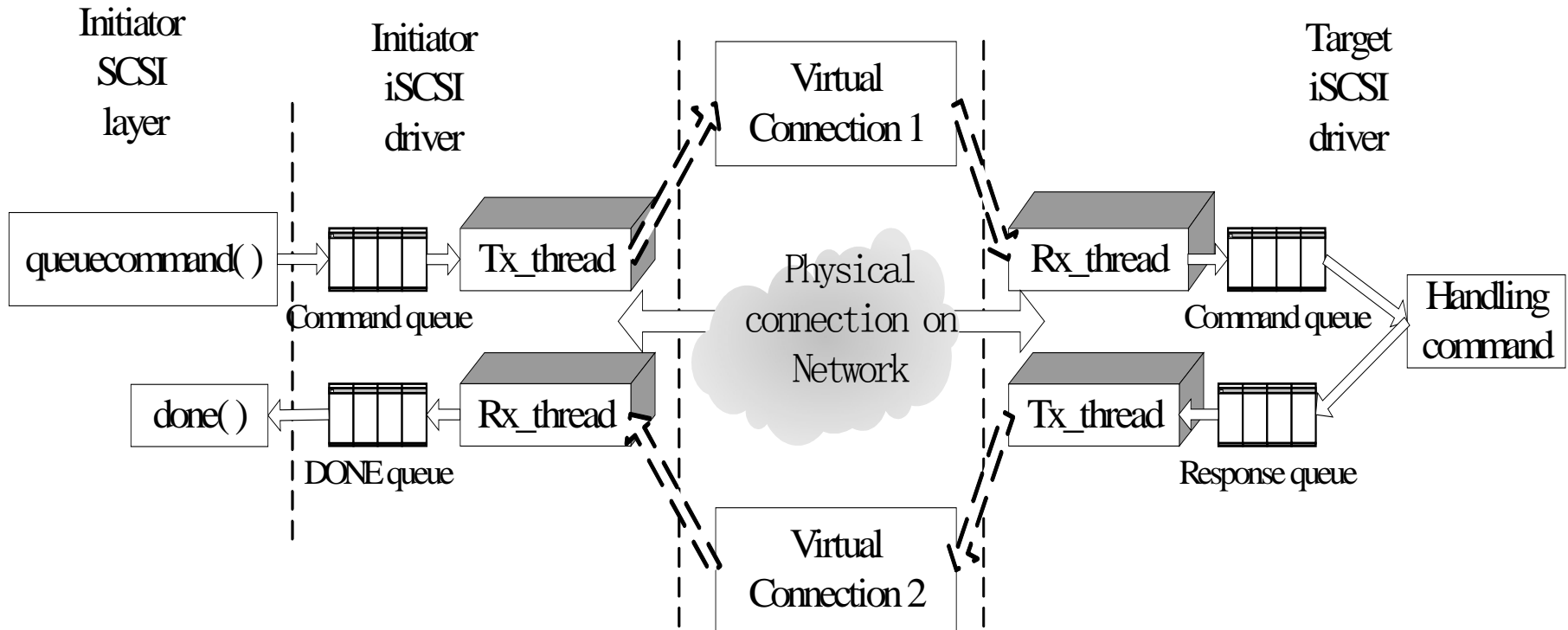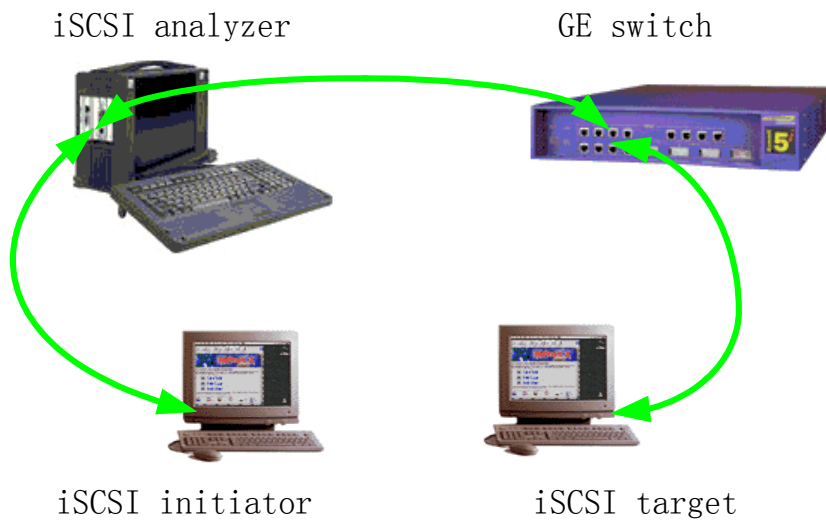- Zero copy
- Local cache

Our Approach

Can we utilized the GE channel bandwidth more effectively in iSCSI?

- Multiple connections

Multiple Connections iSCSI
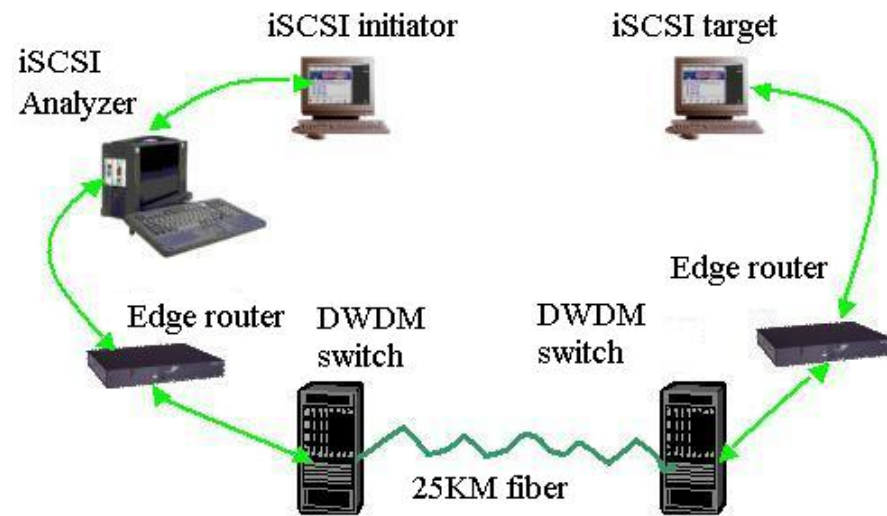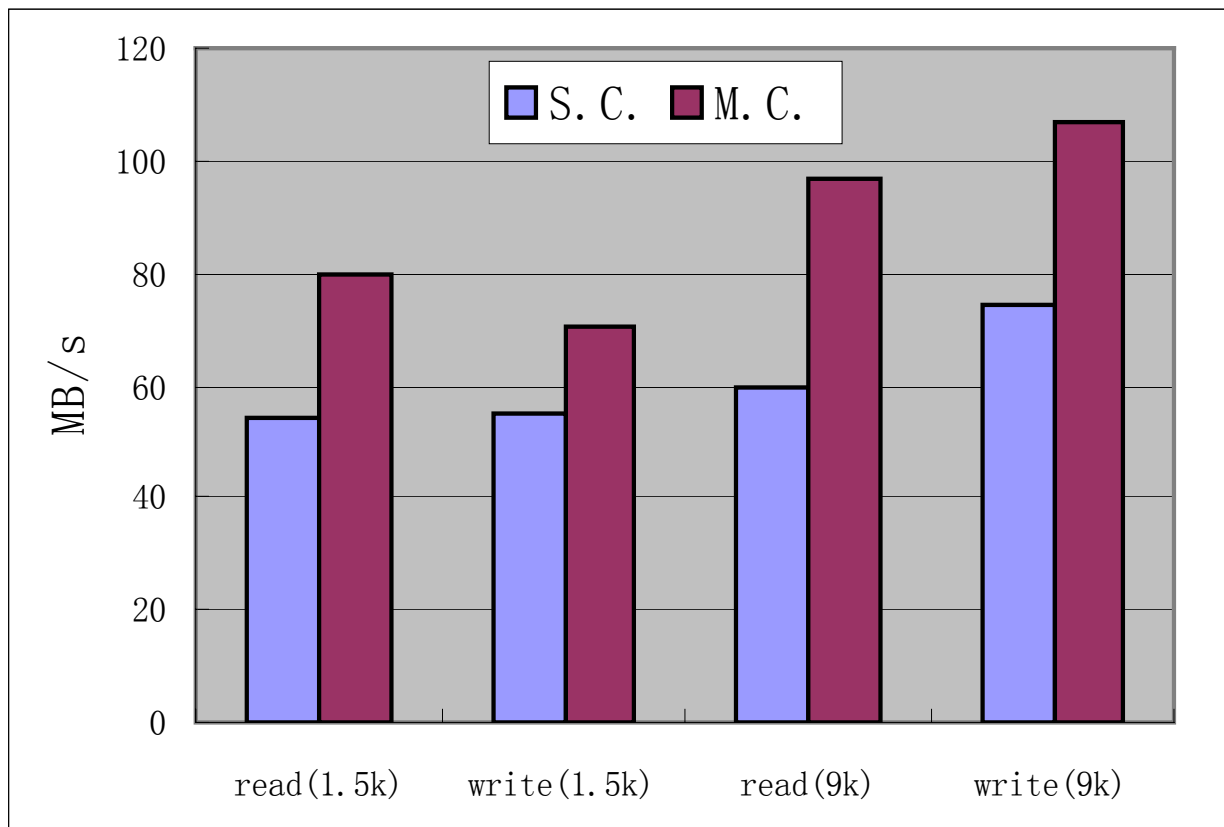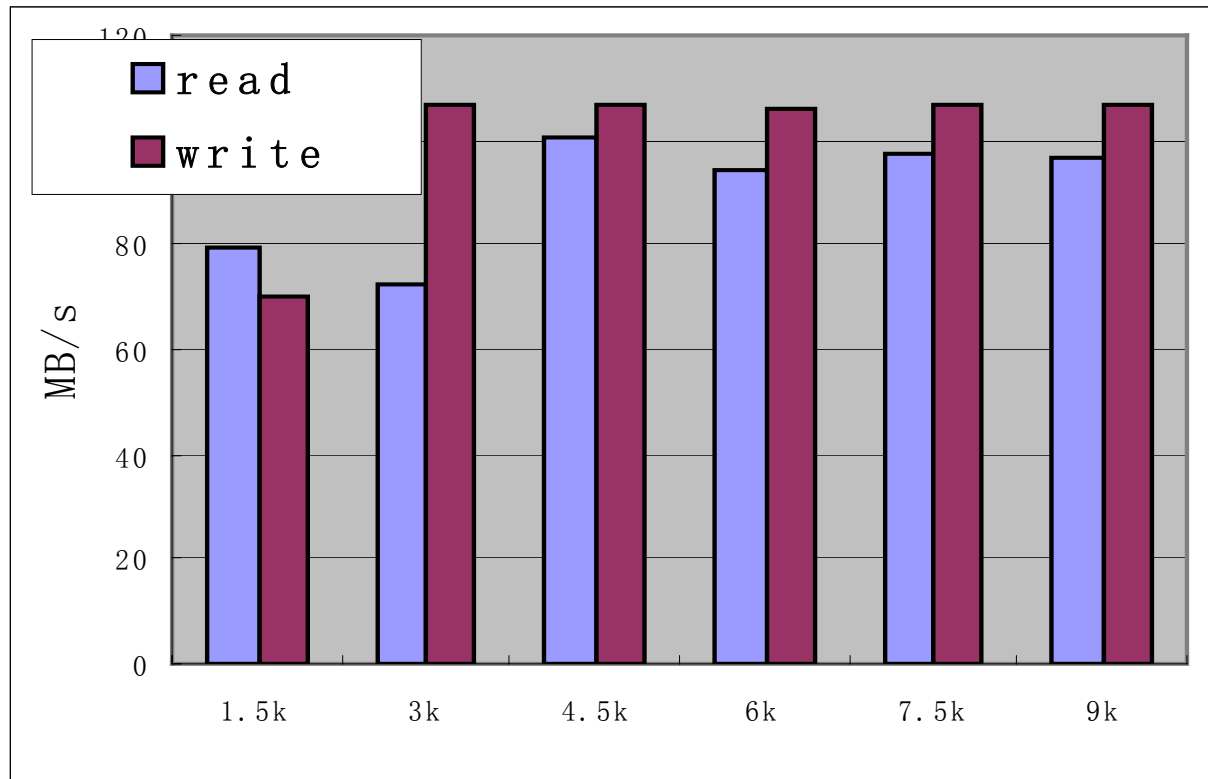
Experiment

Lab Experiment

Metro network experiment
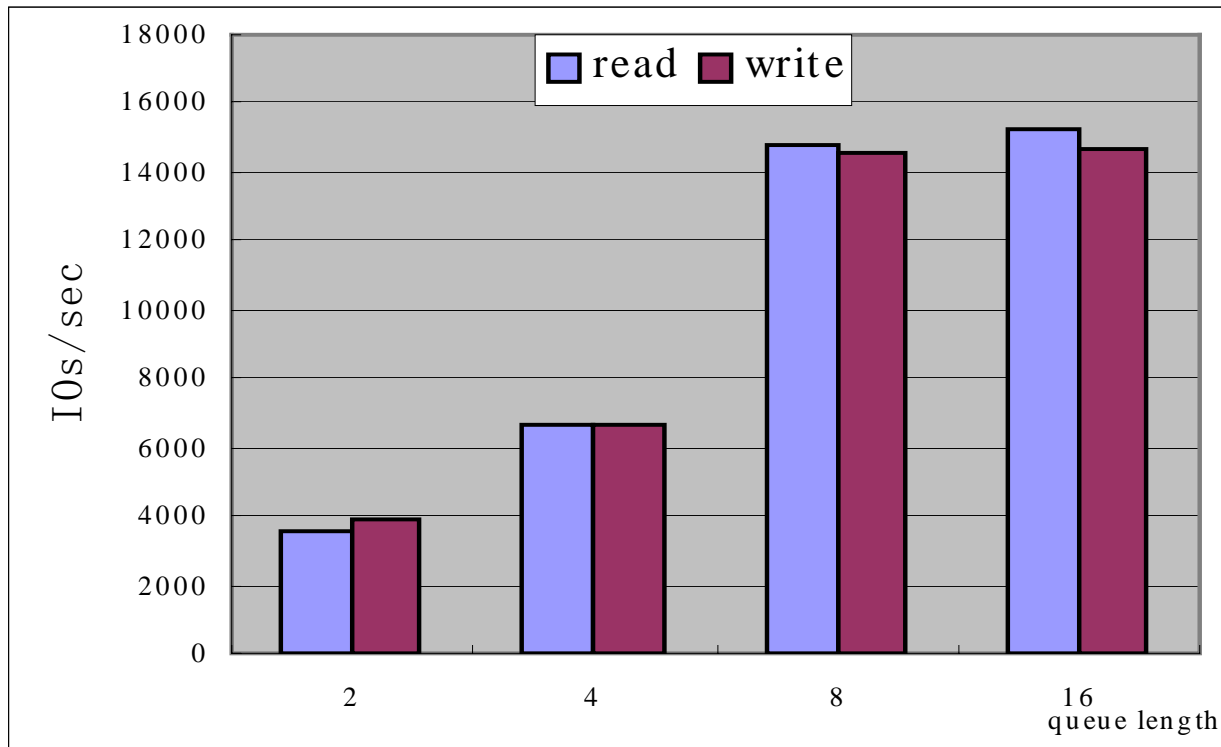
Single vs Multiple Connection

- Multiple connections is better than single connection: 20%~60% improvement

- Jumbo frame is better than normal frame
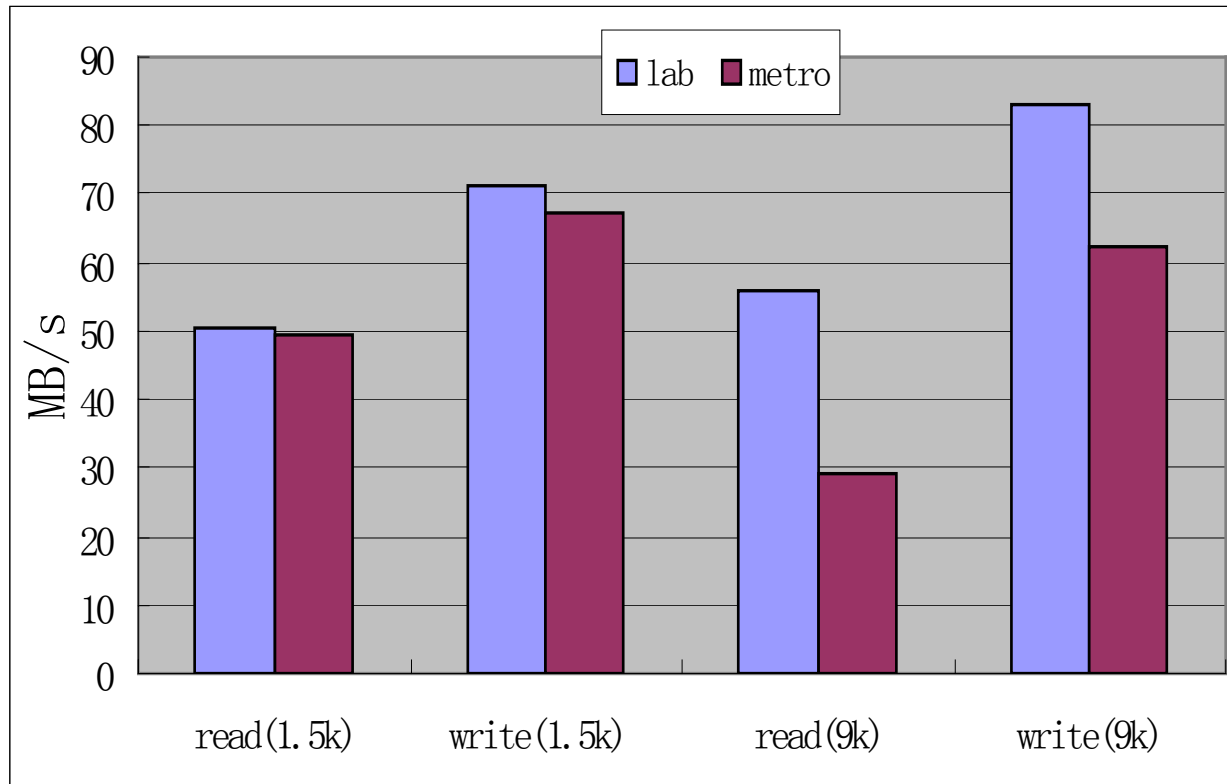
Impact of Frame Size

• Bigger frame will improve performance

Impact of Queue Length on Small IO

- I/O rate increases with the queue length
- Maximum effective queue length is 8

Lab vs Metro Network

• In metro network , throughput is only 2%~5% less than that of lab

Conclusion

- **We have implemented an multiple connection iSCSI prototype**

- **We have conducted experiment both in lab and metro network**

- **The prototype can improve 20%~60% throughput**

- **The prototype can reach 15000 IOPS for small IO**

- **The prototype can work well in metro network environment and cause only ~5% loss of throughput**