

Performance Evaluation of Commodity iSCSI-based Storage Systems

Dimitrios Xinidis, Michail D. Flouris, and
Angelos Bilas

{*dxinid*, flouris, bilas}@ics.forth.gr

**Institute of Computer Science (ICS)
Foundation for Research and Technology – Hellas (FORTH)
Heraklion, Greece**

Motivation

- Storage systems need to scale to large sizes
- Need for providing scalability at the block level
- Connect many disks in the system through scalable network
- Technologies commonly used
 - SCSI
 - Fiber Channel
- Rely on custom network components (Fiber Channels interconnects, controllers)
- Replace custom interconnects with commodity networks even inside the data center
- Main problem is the protocol used on the top of this interconnect

Motivation (cont'd)

- SCSI over IP (iSCSI)
- Network standard that provides a transport layer for SCSI commands over TCP/IP
- Use existing infrastructures for access storage
- IP-based infrastructure has been improving dramatically
 - 10Gbit Ethernet will become commodity over the next 2 years
- It is not clear what is its impact on system performance
- Previous work has revealed the adverse effect of TCP processing
- How to map iSCSI related overhead to different layers in I/O protocol stack of the kernel?

Goals

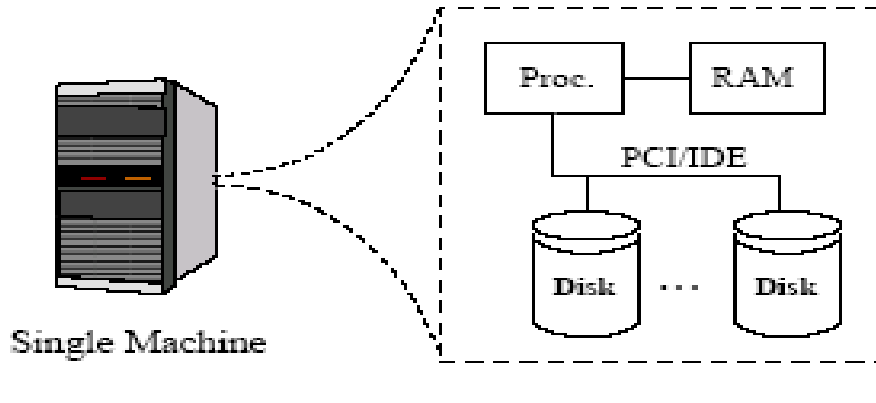
- Examine the contribution of each component of the I/O protocol stack in overhead in iSCSI based systems
- Impact of iSCSI to application server performance
- Use both application and micro benchmarks

Platform

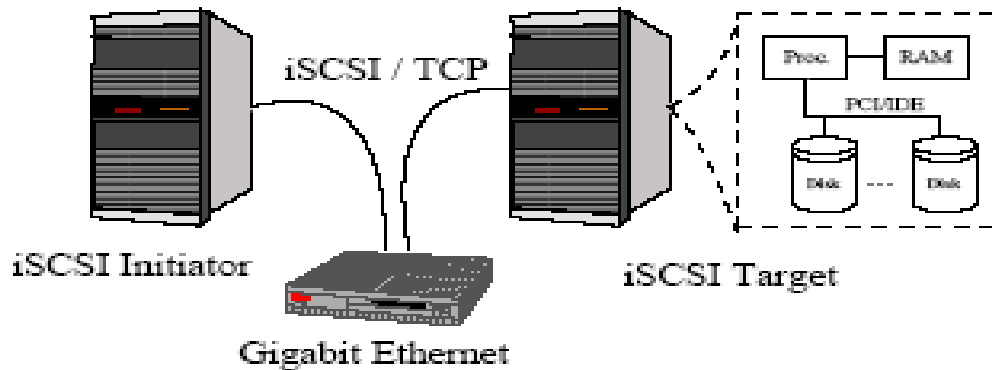
- 16 dual-processors, Athlon at 1.8 GHz, 512 MB RAM
- Each node has 1 Gbps network interface
- Nodes are connected with 24-port switch, 48Gbit/s backplane
- Nodes are divided in application and storage nodes
- Storage nodes have 6 disks, 80GB – 2MB cache
- 32 bit/33 MHz PCI bus, peak throughput 133MBytes/s
- Disks are configured in RAID-0 mode with Linux MD driver
- Intel iSCSI implementation
- Linux Redhat9.0, kernel 2.4.x

Setups

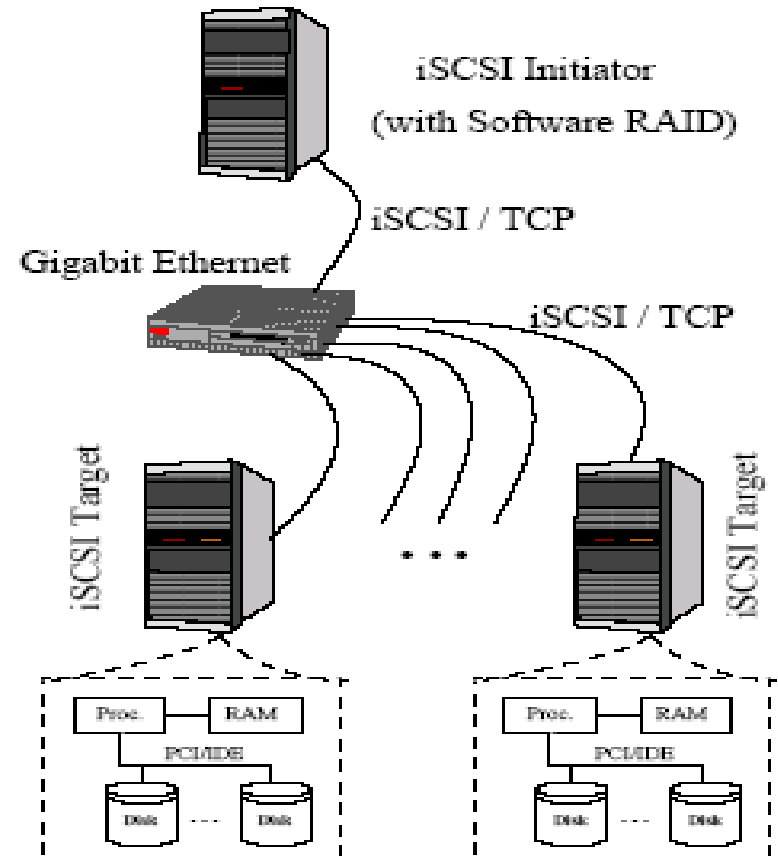
(A) Direct-attached Disk Setup



(B) iSCSIx1 Setup

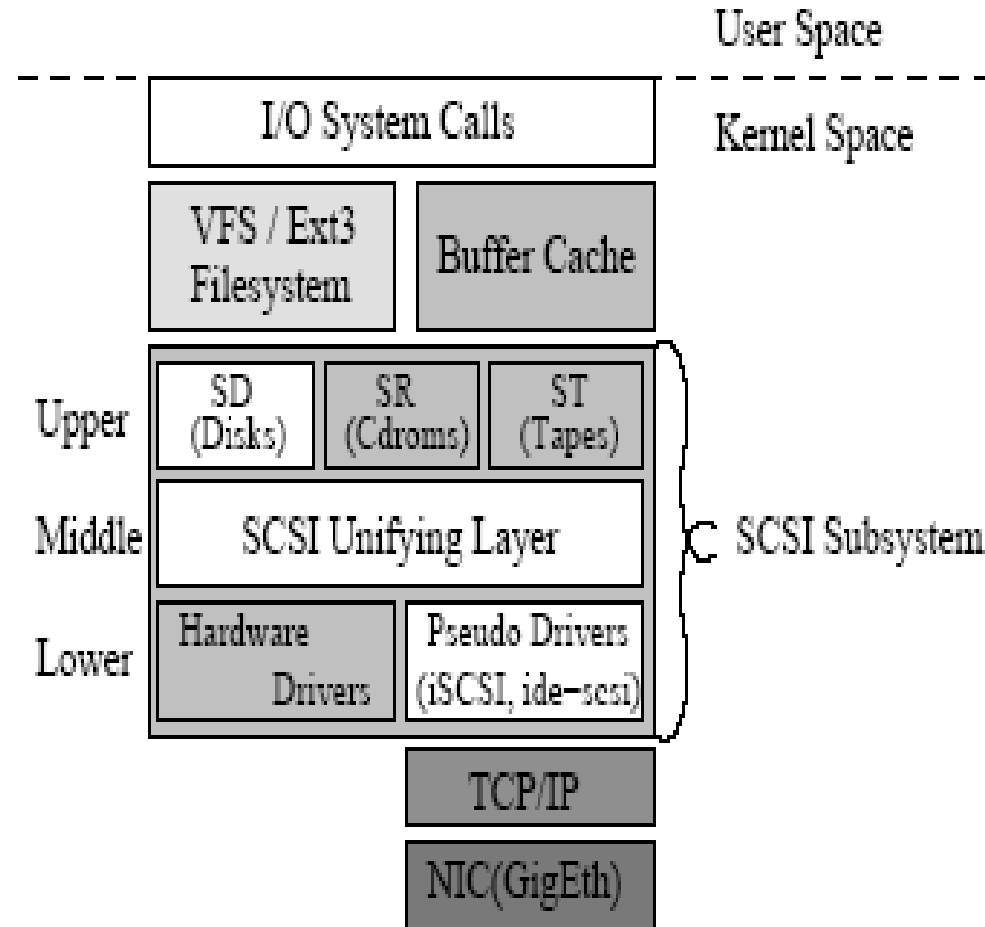


(C) iSCSIx3 Setup



Methodology

- What is the cost of each of these layers?
- Instrument kernel layers
- Develop time framework presented at CAECW 05 workshop
- Instrumentation is a highly time consuming procedure



Benchmarks

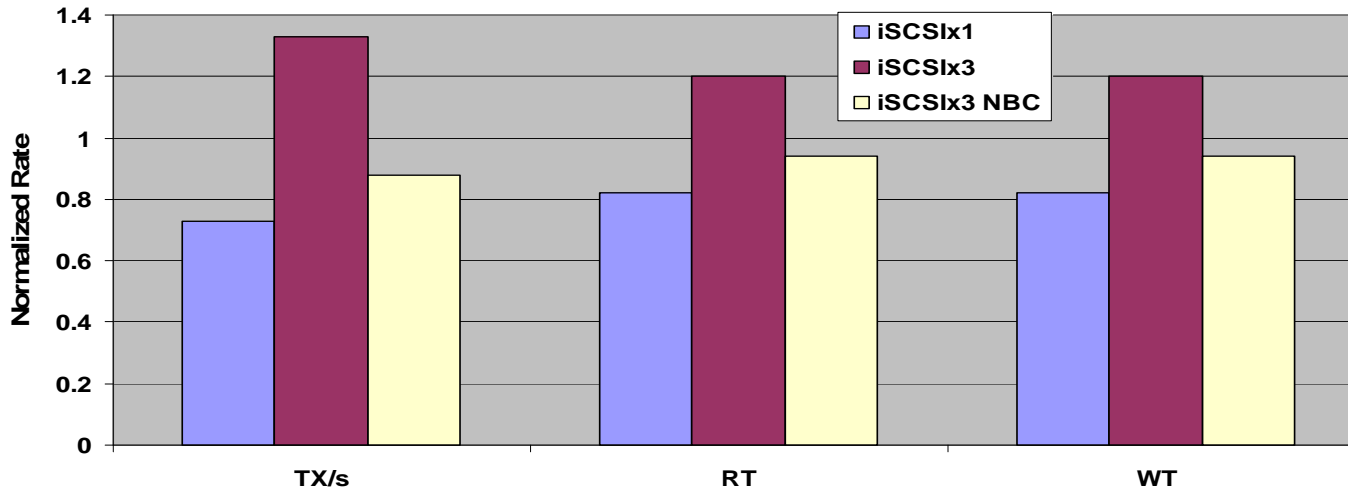
- Use both real applications and micro benchmarks
- Micro benchmarks
 - Iometer with workloads variations, where we vary access patterns and read-write mixes
 - Postmark, simulates the behavior of an Internet Mail Server
- Real applications
 - MySQL, subset of TPCH workload
 - SpecSFS: Measures the throughput and response time of an NFS server

Base Performance

Configuration	Read (Mbytes/s)	Write (Mbytes/s)
Single disk	50	20
Direct Attached	120	50
iSCSI	25	40
iSCSIx3	50	40

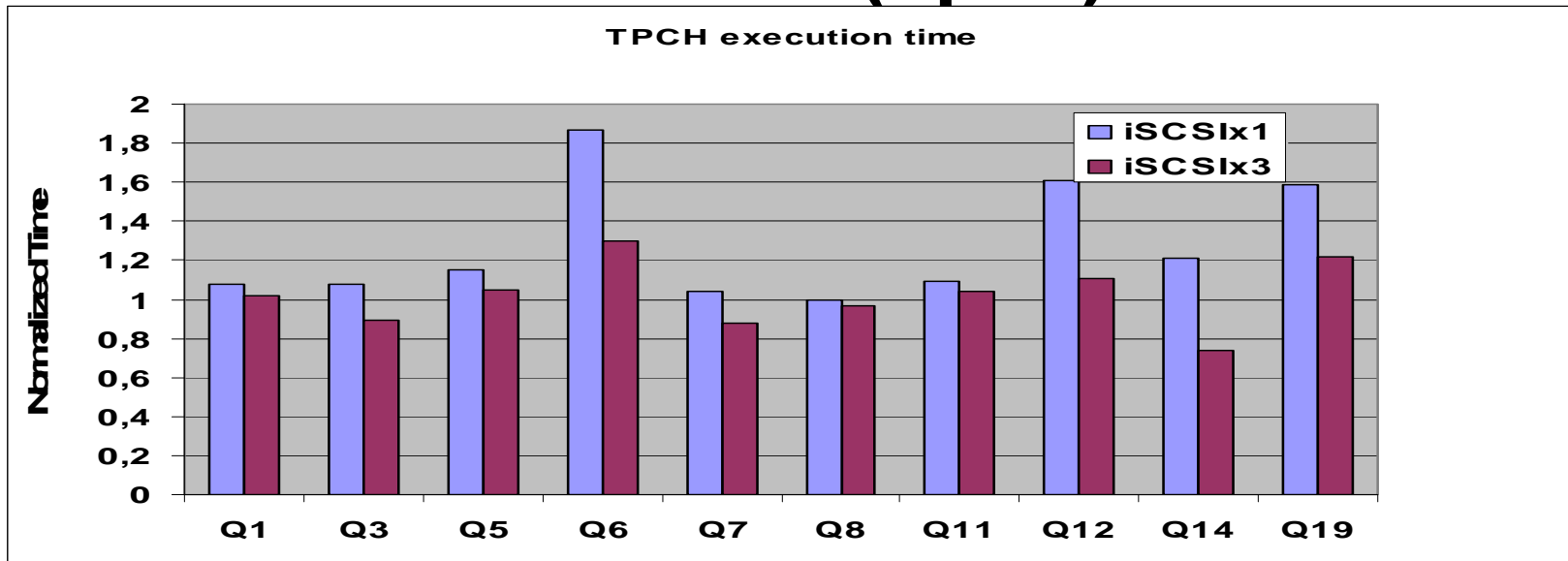
- Iometer with sequential reads and writes
- Request size is 4, 8 Kbytes
- Maximum throughput in direct configuration is the maximum of PCI bus

Results (Postmark)



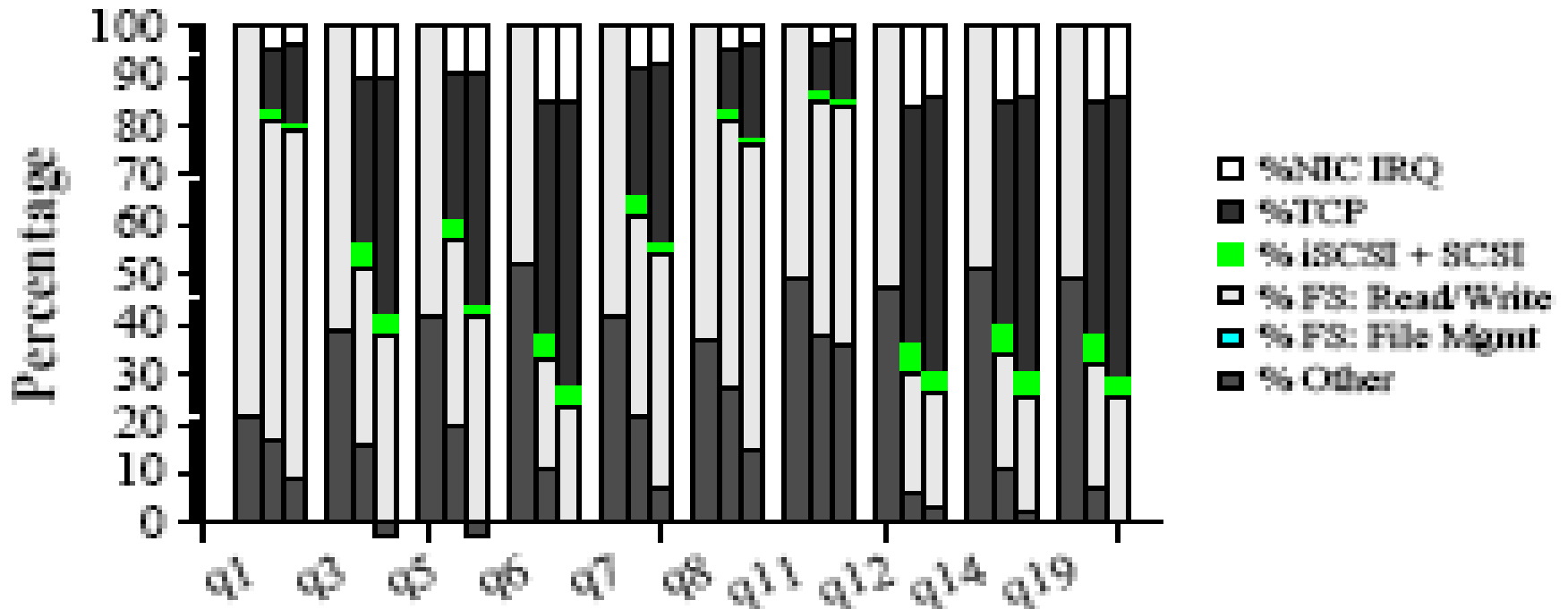
- Sensitive to I/O latency
- iSCSI reduces performance up to 30%
- iSCSIx3 improves performance up to 45%
- This benefit comes from the additional cache and not from having more disks

Results (Tpch)



- TPCH mostly depends on I/O throughput
- iSCSI reduces performance up to 86% and 26% on average
- iSCSIx3 performance improves significantly
- Tpch benefits mostly from the increased number of disks under iSCSIx3 (Not from increased buffer cache)

System Time Breakdown



- Besides TCP an important component of the time is spent in buffer cache and system call interface

Conclusions

- Examine the impact of using commodity interconnects and protocols to scale storage systems
- iSCSI has a significant impact in all applications we have examined
- iSCSIx3 is able to scale system resources and recover most of the performance loss
 - Postmark benefits from the increased buffer cache, TPCH benefits from the increased number of disks
- Examine detailed breakdown of I/O overheads to layers of the I/O protocol stack
 - Besides TCP and NIQ, buffer cache management and system call interface is a significant component