Performance Evaluation of iSCSI over Local, Metropolitan, and Wide Area Networks

Ismail Dalgic
Director of Architecture
Intransa, Inc.
ismail.dalgic@intransa.com

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
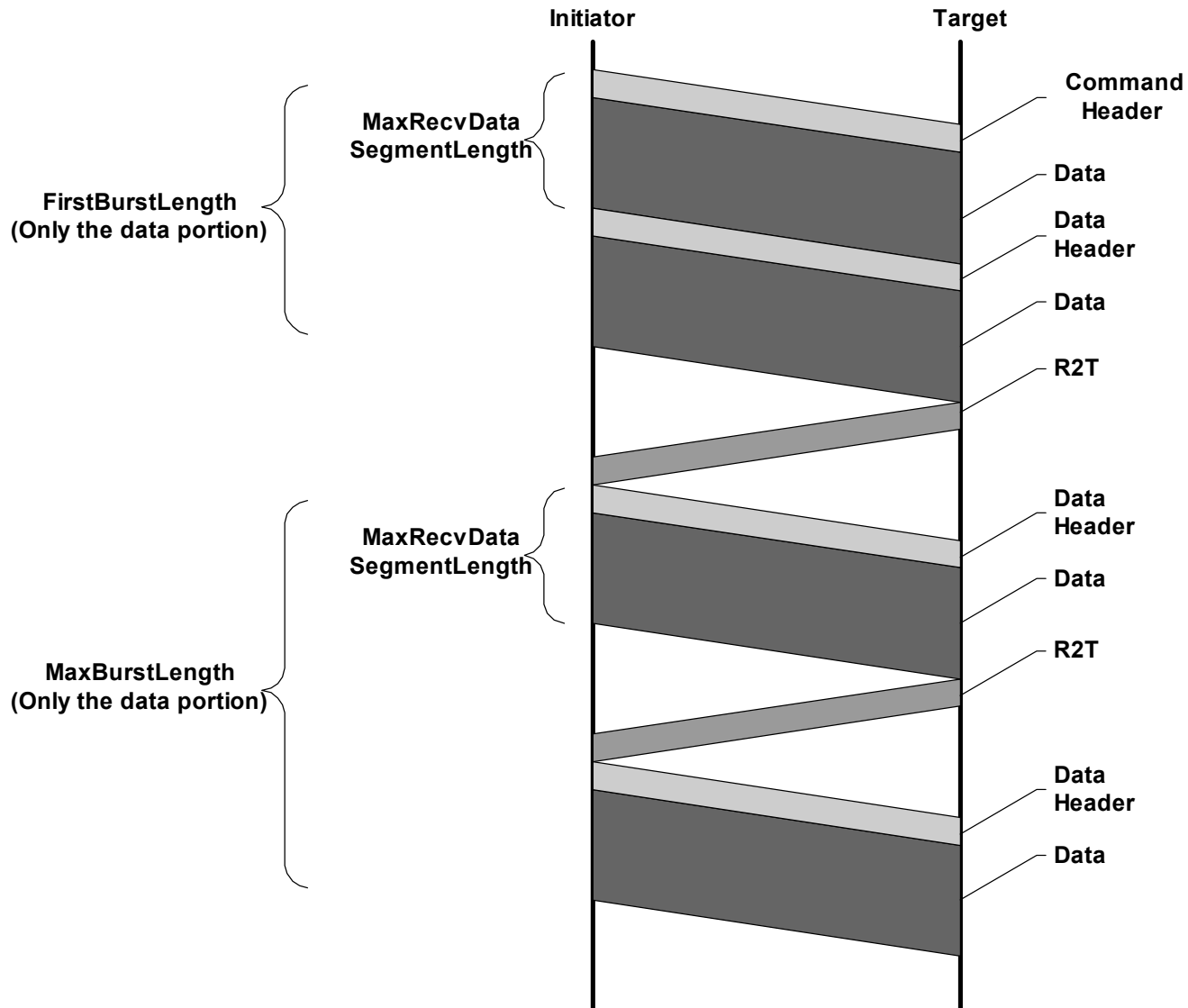Adelphi MD USA
April 13-16, 2004

Objectives

- Investigate the effect of latency on iSCSI performance

- Identify bottlenecks and the tunable parameters to eliminate them

- Adjust tunable parameters to improve performance

Tunable Parameters

- iSCSI level
  - Command request size
  - Command window credit amount
  - Sending solicited vs. unsolicited data (for writes)
  - Number of simultaneous iSCSI connections per session
- TCP level
  - Transmit and receive window sizes
- Network / Data Link level
  - Path MTU
  - Transmit and receive buffer sizes

Solicited vs. Unsolicited Writes

Initiator

Target

MaxRecvData SegmentLength

FirstBurstLength (Only the data portion)

MaxRecvData SegmentLength

MaxBurstLength (Only the data portion)

Command Header

Data

Data Header

Data

R2T

Data Header

Data

R2T

Data Header

Data

Write Scenarios Considered

- Most Solicited
  - *FirstBurstLength = 0*
- Most Unsolicited
  - *FirstBurstLength >= max write-request size*

Experimental Setup

iSCSI Initiator

LANForge WAN Emulator
set at OC-3 rate (155 Mbits/s)

SC5100
Storage
Controller

Intransa
IP5000
Storage
System

GE Switch

DE5200
Disk
Enclosure
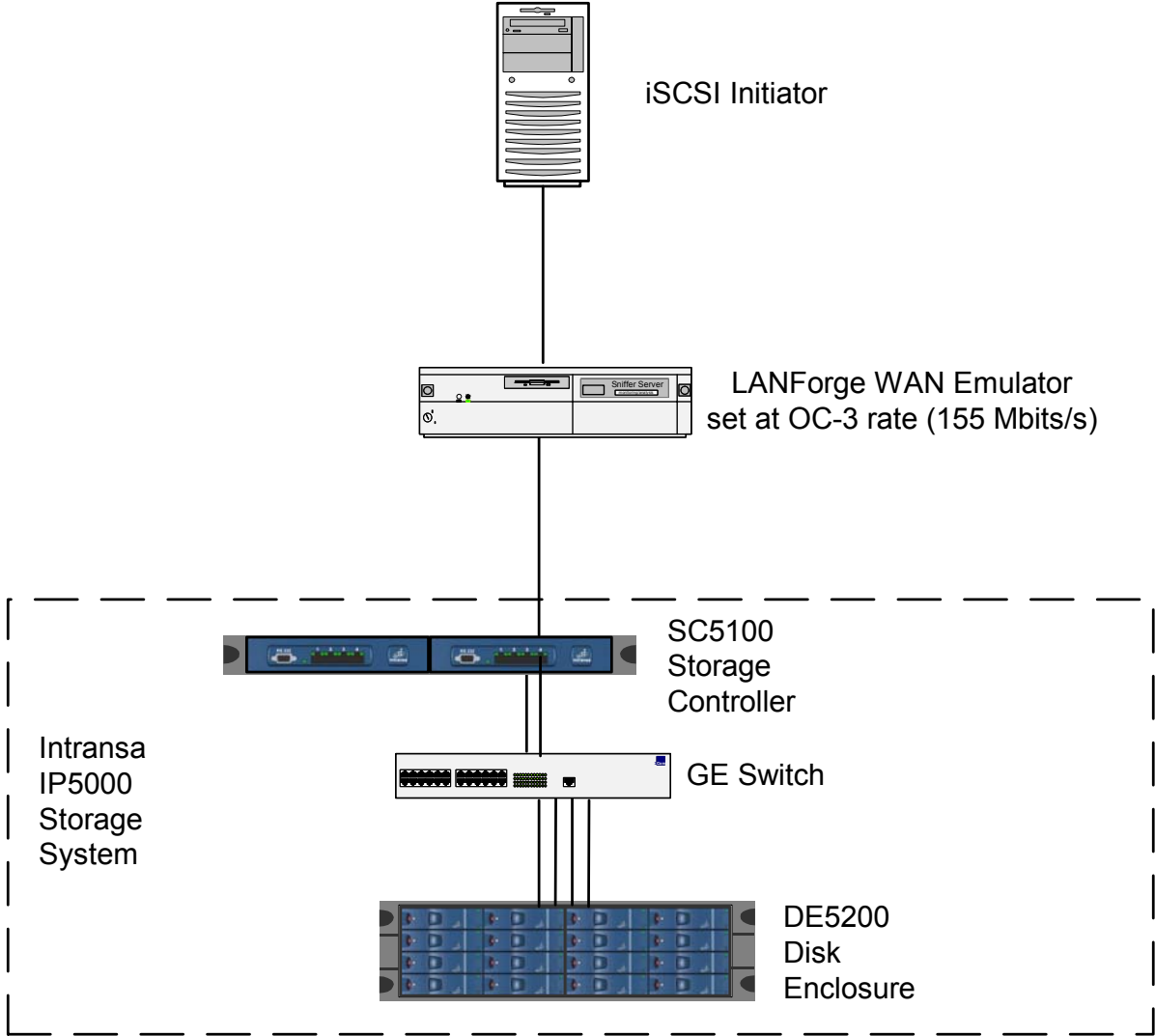
TCP Throughput with Default Window (64 Kbytes)

Transfer Direction	Round Trip Latency			
	0 ms	2 ms	10 ms	50 ms
I → T	19.0	18.7	4.4	0.9
T → I	19.0	13.5	3.2	0.7

Throughput in MB/s

TCP Throughput with Large Window (10 MBytes)

Transfer Direction	Round Trip Latency			
	0 ms	2 ms	10 ms	50 ms
I → T	19.3	19.3	19.3	19.3
T → I	19.3	19.3	19.3	19.3

Throughput in MB/s

In the remainder of this study, we use this TCP window size

iSCSI Throughput: Most Solicited Writes

Request Size	No. of outstanding iSCSI commands	Round Trip Latency			
		0 ms	2 ms	10 ms	50 ms
1 kB	1	1.3	0.2	0.05	0.01
	32	11.3	3.3	0.8	0.2
8 kB	1	7.7	1.7	0.4	0.08
	32	19.0	18.9	5.8	1.2
64 kB	1	19.2	10.9	2.9	0.6
	32	19.2	19.3	19.3	4.5
256 kB	1	19.2	19.2	7.8	1.7
	32	19.3	19.3	19.3	4.7

Throughput in MB/s

iSCSI Throughput: Most Unsolicited Writes

Request Size	No. of outstanding iSCSI commands	Round Trip Latency			
		0 ms	2 ms	10 ms	50 ms
1 kB	1	2.2	0.4	0.09	0.02
	32	17.4	6.6	1.5	0.3
8 kB	1	10.3	3.2	0.7	0.2
	32	19.2	19.2	11.6	2.5
64 kB	1	19.3	17.6	5.4	1.2
	32	19.3	19.3	19.3	4.8
256 kB	1	19.3	19.3	11.6	2.5
	32	19.3	19.3	19.3	4.9

Throughput in MB/s

iSCSI Throughput: Reads

Request Size	No. of outstanding iSCSI commands	Round Trip Latency			
		0 ms	2 ms	10 ms	50 ms
1 kB	1	2.3	0.4	0.1	0.02
	32	17.6	8.0	2.3	0.5
8 kB	1	10.2	3.1	0.8	0.2
	32	19.2	19.2	19.1	4.7
64 kB	1	19.2	17.2	5.4	1.2
	32	19.2	19.3	19.3	19.2
256 kB	1	19.2	19.2	18.3	4.7
	32	19.3	19.3	19.3	19.3

Throughput in MB/s

iSCSI Throughput: Gigabit Ethernet LAN

Request Size	No. of outstanding iSCSI commands	Most Solicited Writes	Most Unsolicited Writes	Reads
1 kB	1	3.4	5.5	5.4
	32	17.1	23.4	19.9
8 kB	1	17.4	25.0	22.5
	32	68.6	84.3	72.3
64 kB	1	56.1	65.9	61.3
	32	96.5	99.8	91.9
256 kB	1	71.3	79.9	74.7
	32	97.3	100.4	98.4

Throughput in MB/s

Conclusions

- Default TCP window size is inadequate for good WAN or MAN performance.
- Product of the iSCSI command window and request size has a significant impact on throughput
- Most solicited writes result in a major performance penalty
- With adequate tuning, iSCSI over TCP is capable of achieving good throughput in all kinds of networks