

iSCSI Testing At NASA GSFC

Hoot Thompson

Patuxent Technology Partners, LLC

11030 Clara Barton Drive

Fairfax Station, VA 22039

Phone: +1-703-250-3754 FAX: +1-703-250-3742

e-mail: hoot@ptpnow.com

Presented at the THIC Meeting at the

Raytheon ITS Auditorium,

1616 McCormick Dr., Upper Marlboro MD 20774-5301

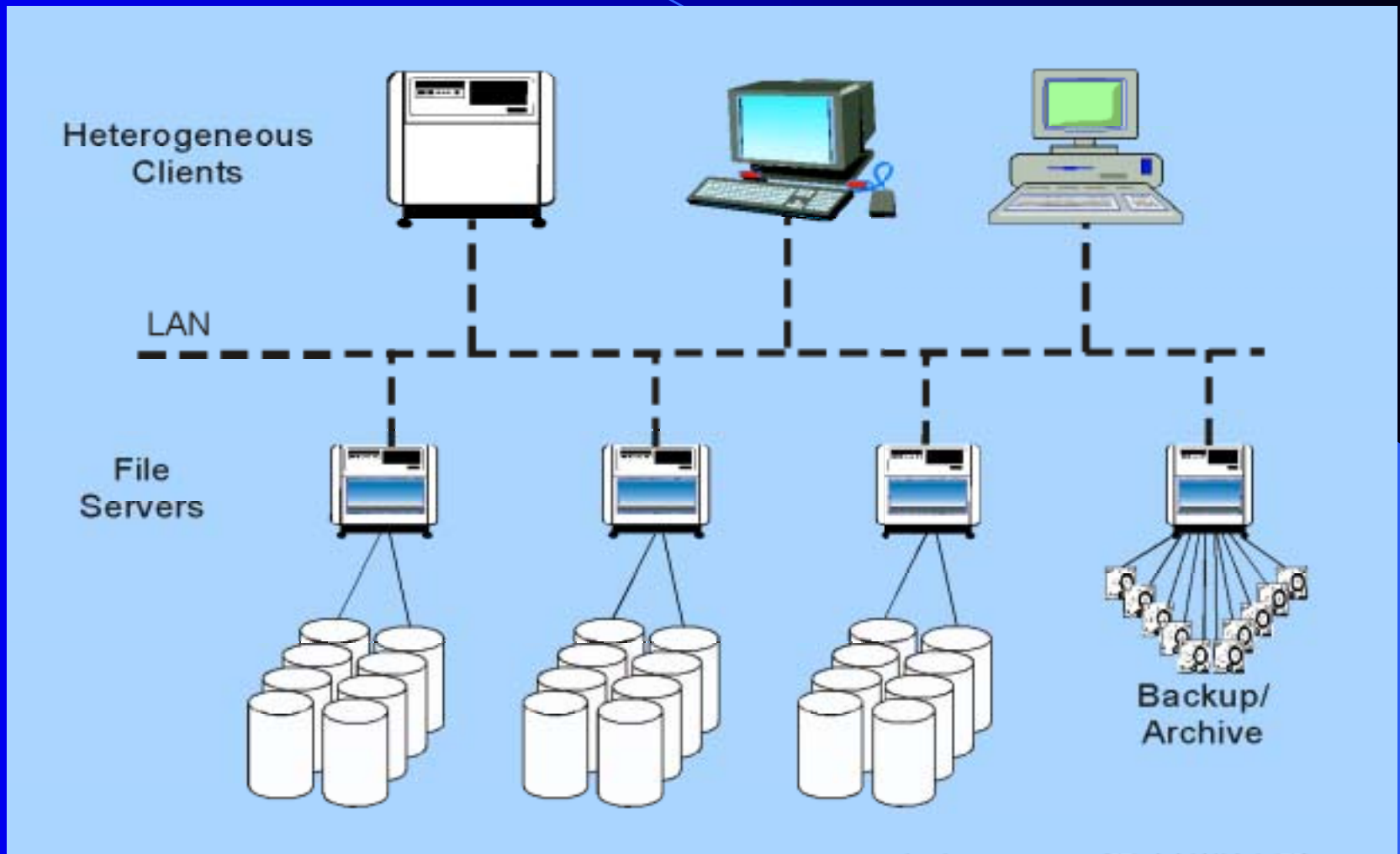
November 5-6, 2002

Introduction – Some Trends

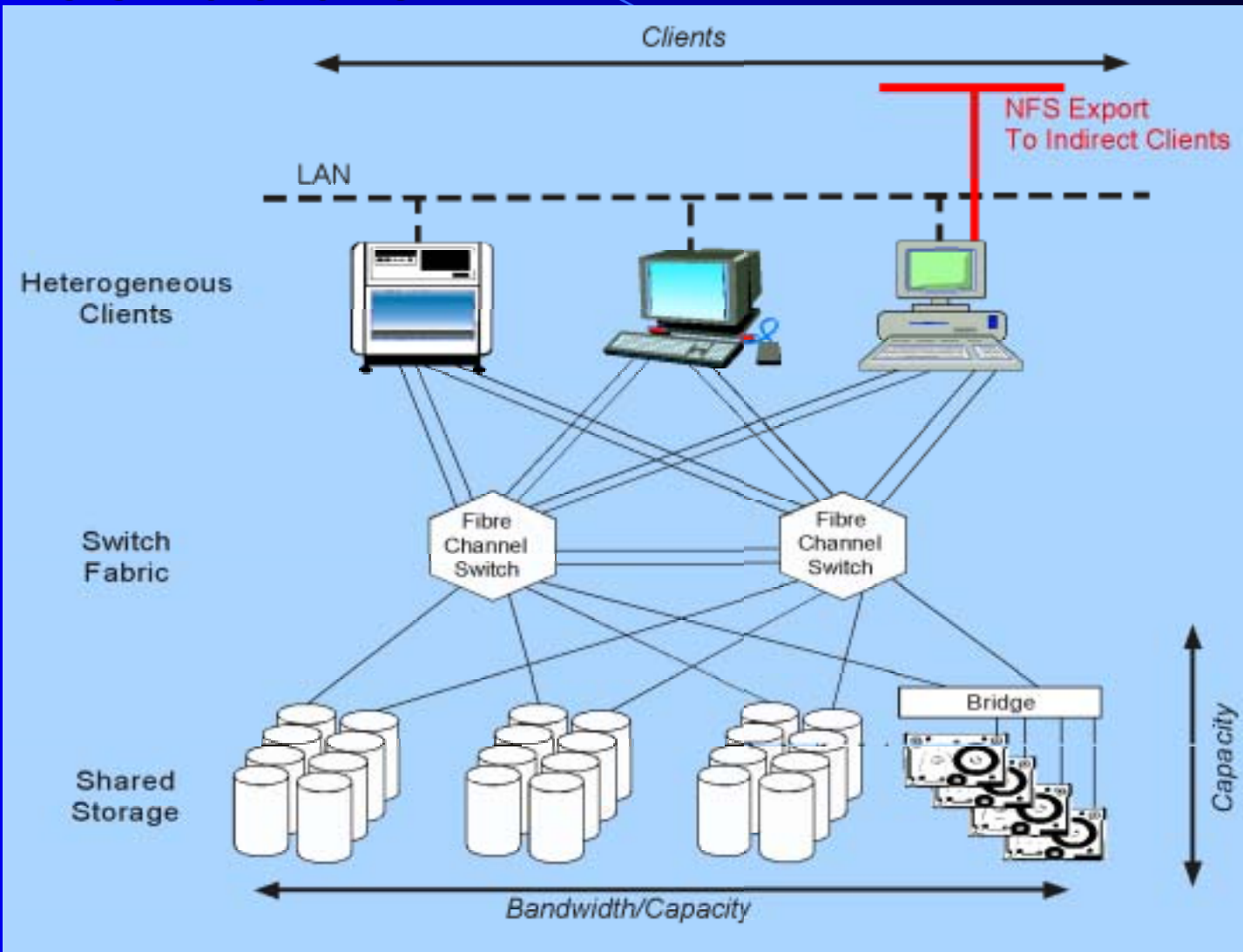
- Storage consolidation continues to be a popular architectural construct – Storage Area Networks (SAN)
- Minimizing costs while expanding storage capacity and access is an obvious driver
- Fibre channel connected storage remains dominant at least in the high performance space
- Existing IP infrastructures versus ‘still’ expensive fibre channel represent an attractive mechanism for connecting users to storage at the block level

Are IP based transport technologies the answer?

By Way of Review – Traditional Infrastructure



Storage Area Network (SAN) Infrastructure



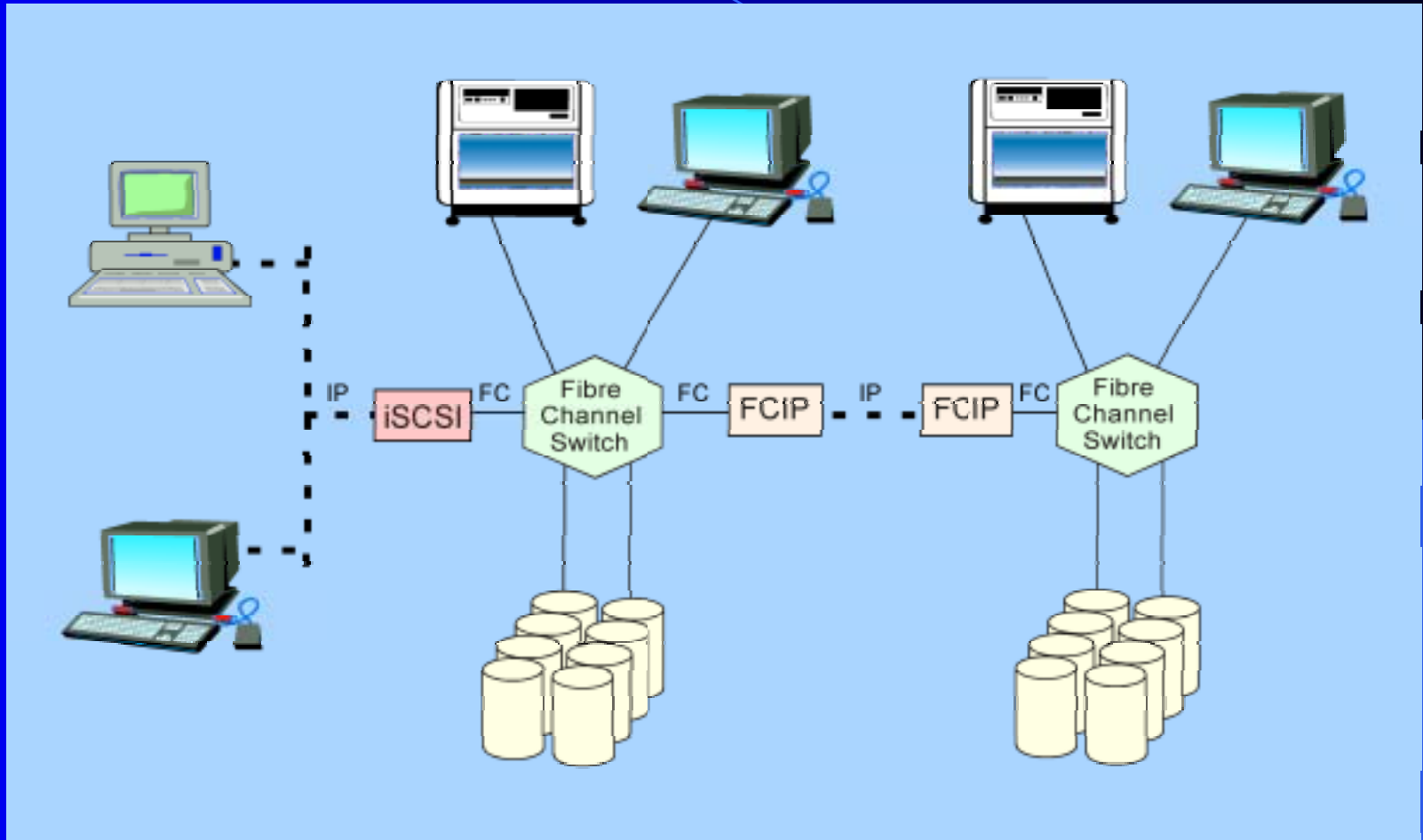
Definitions

- Fibre Channel
 - Industry standard high-speed SCSI transport technology
 - 1 or 2 Gbit/sec, 10 Gbit/sec coming
- Internet SCSI (iSCSI)
 - *“represents a light switch approach to storage networking”*
- Fibre Channel Over IP (FCIP)
 - *“means of encapsulating Fibre Channel frames within TCP/IP specifically for linking Fibre Channel SANs over wide areas”*
- Internet Fibre Channel Protocol (iFCP)
 - *“gateway-to-gateway protocol for providing fibre channel fabric services to fibre channel end devices over a TCP/IP network”*

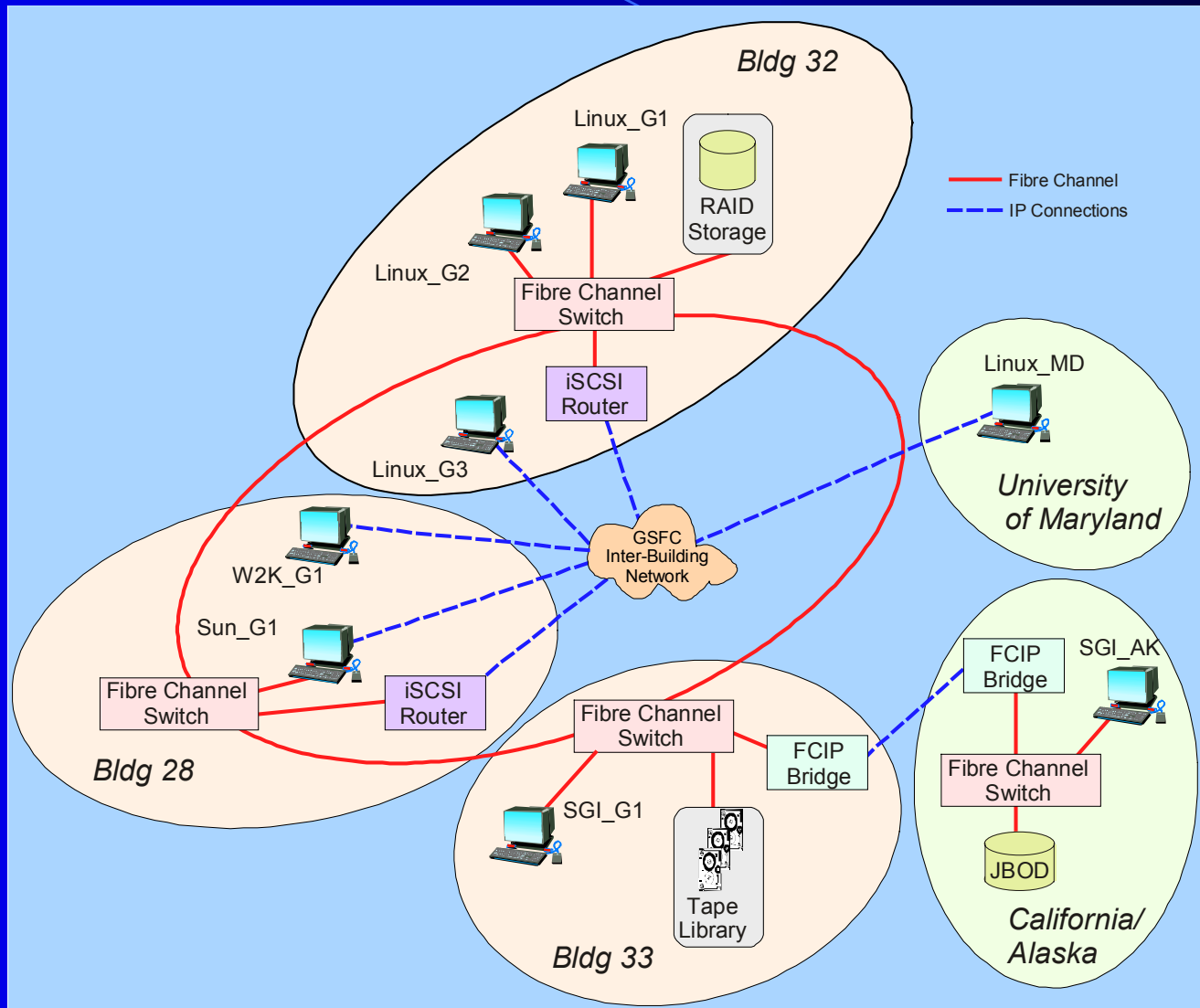
Notes:

- Standards largely the work of Internet Engineering Task Force (IETF) – www.ietf.org
- Definitions from “IP SANs – A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks” by Tom Clark

IP-Based SAN Connections



GSFC Pilot SAN Configuration



IP Testing

- iSCSI in comparison to native fibre channel
- Primarily Linux based machines
 - Sun platform also ‘looked at’
 - SGI currently does not support iSCSI
 - Windows® not the platform of choice
- TCP off-load engine (TOE) card
 - Evaluated on Windows machine

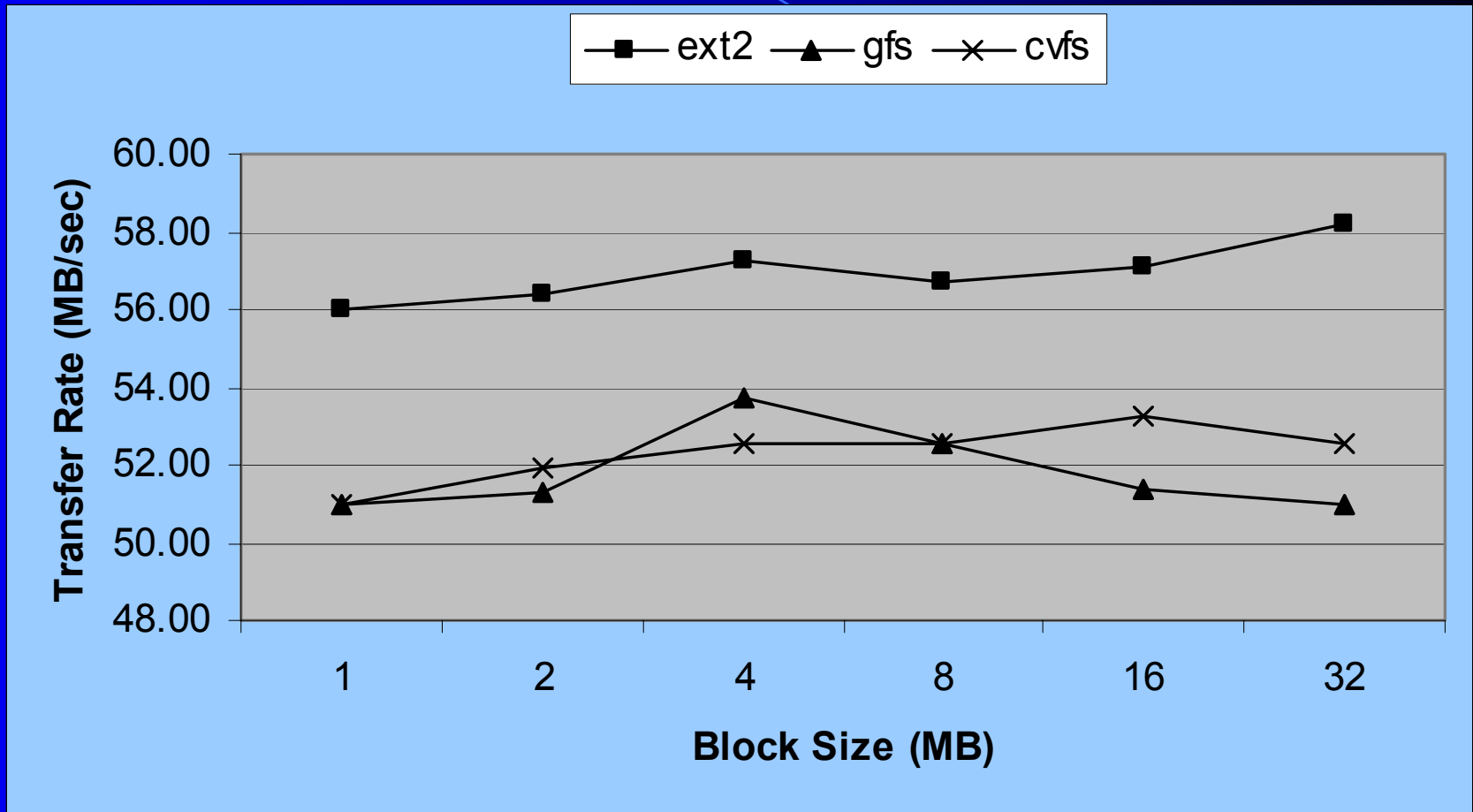
Benchmarks

- Large file, large block sequential transfers
 - *lmd*
- Small file, transaction oriented tests
 - *bonnie++*
 - *Postmark*
- Application testing
 - Composite generation from MODIS data by U of MD staff
- Different file systems
 - ext2fs (native Linux file system)
 - cvfs now StorNext File System (ADIC)
 - Global File System (Sistina)

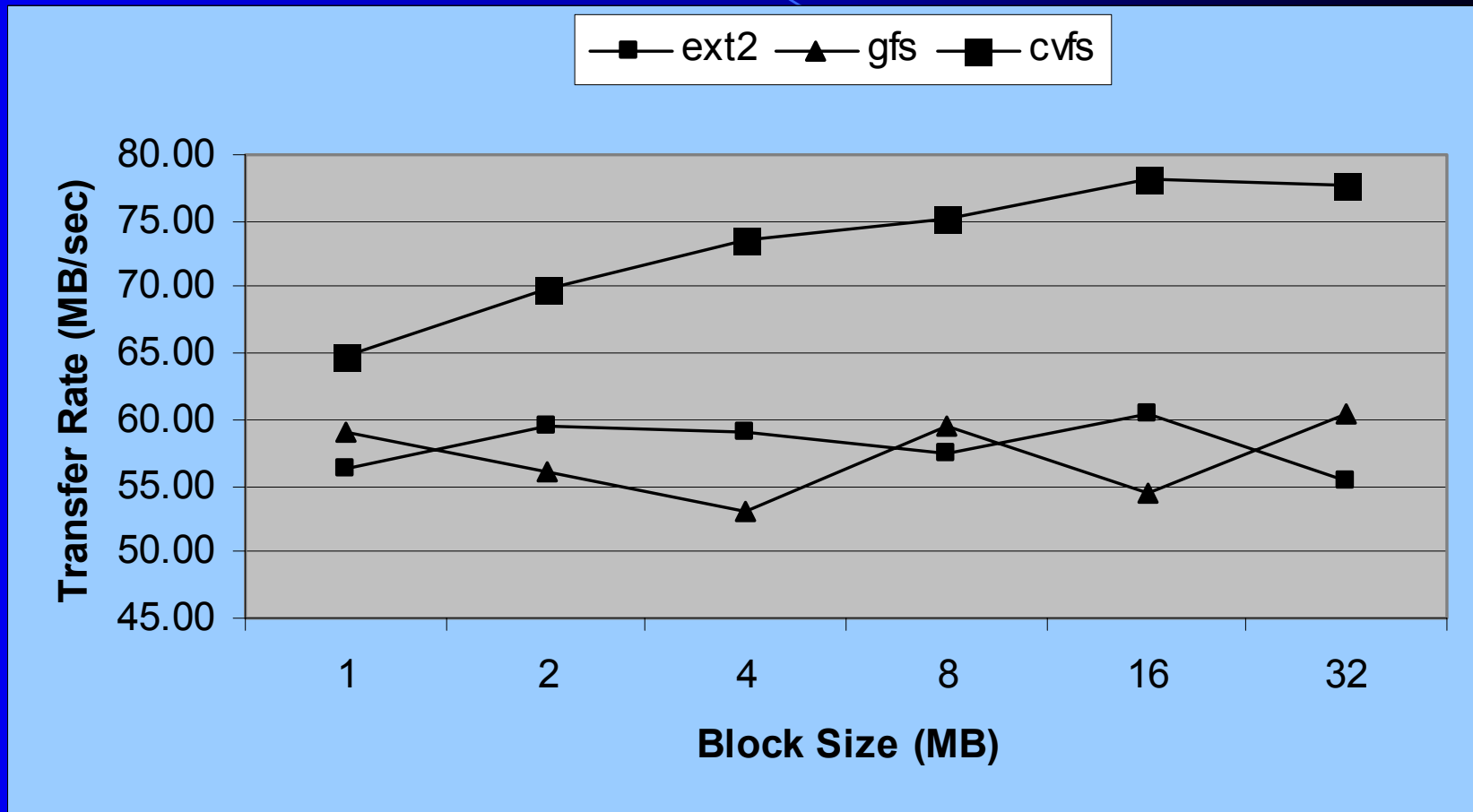
Notes:

- Benchmark numbers for the most part are ‘out of the box’ results and should be viewed as representative not definitive.

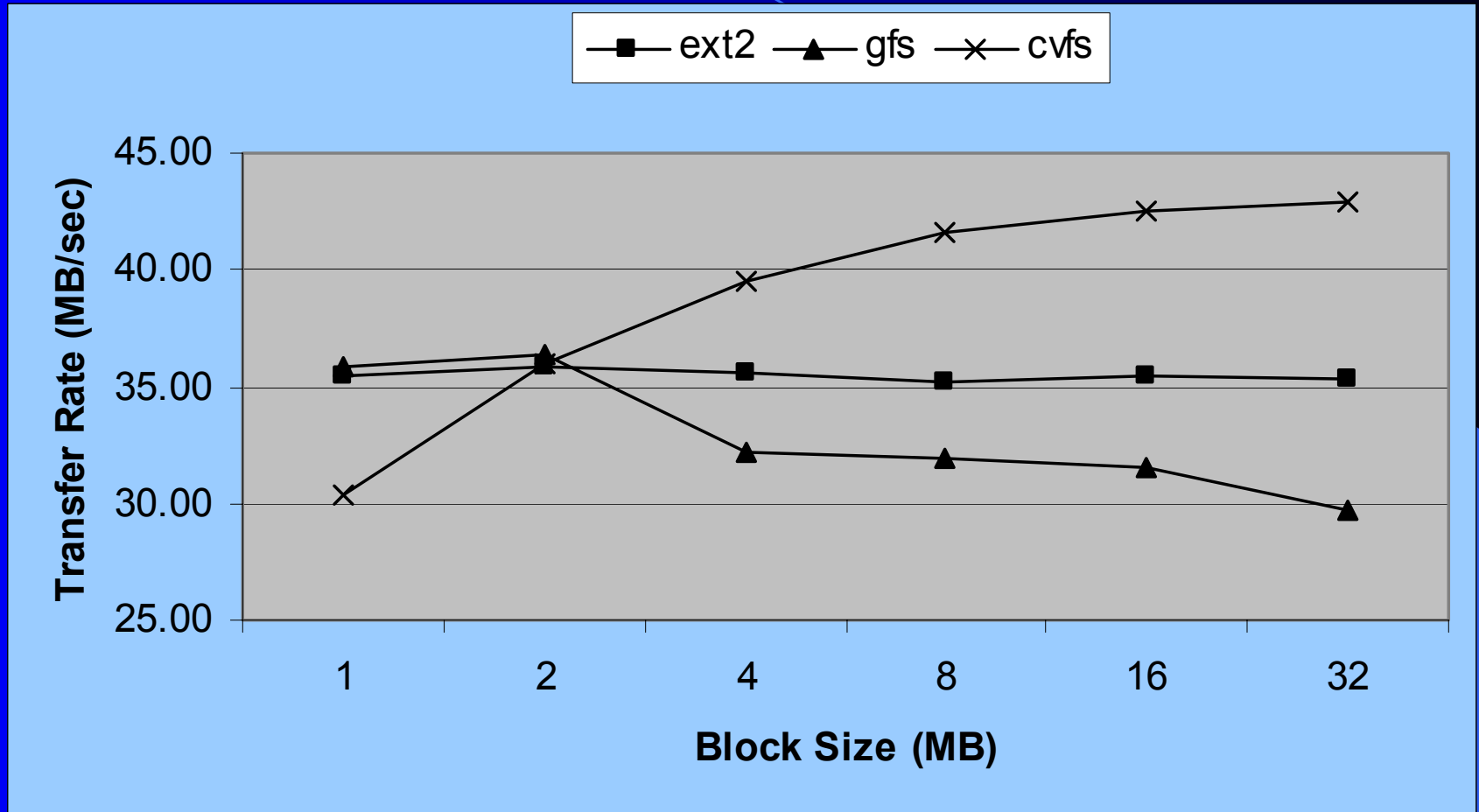
'Linux_G2' Write Performance



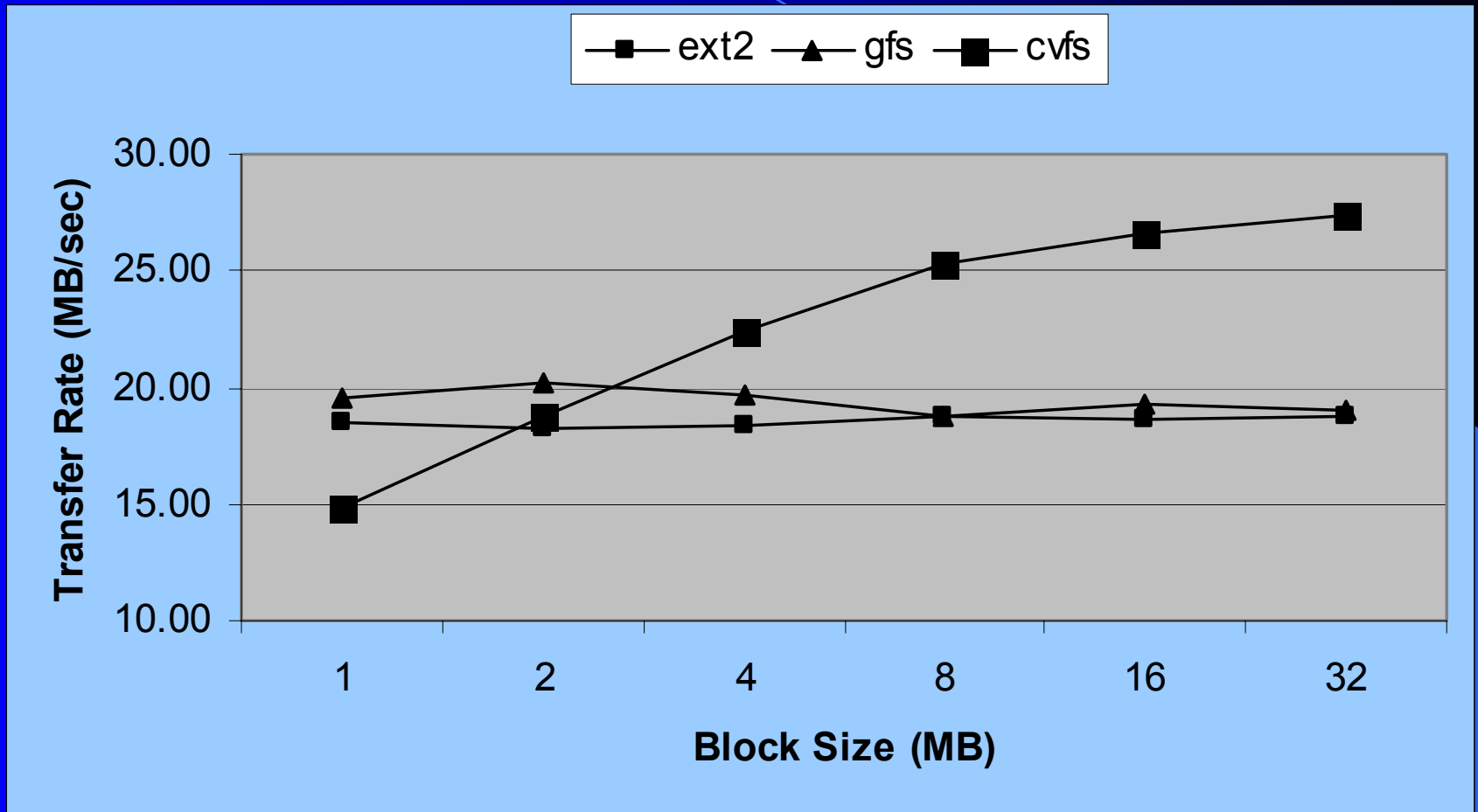
'Linux_G2' Read Performance



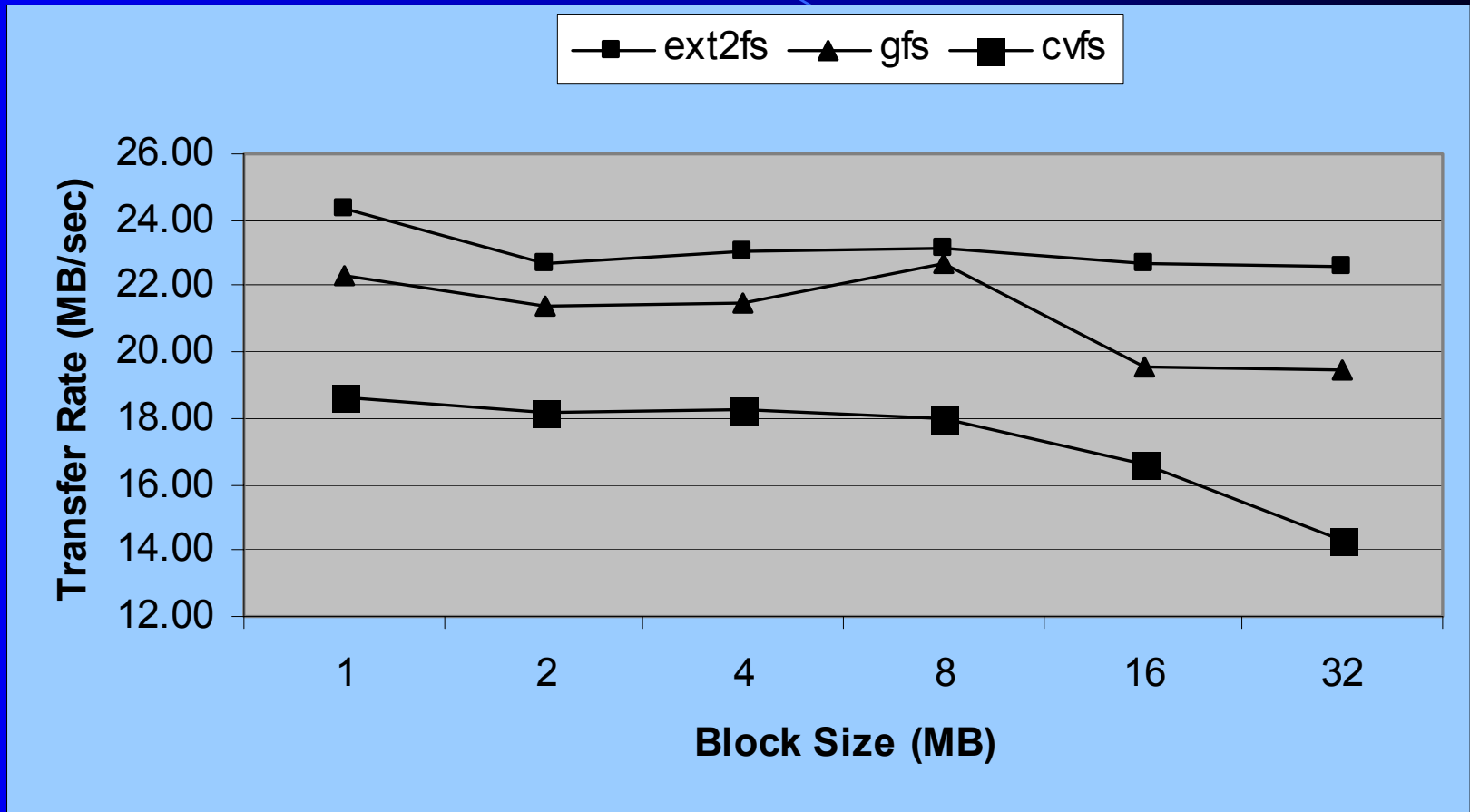
'Linux_G3' Write Performance



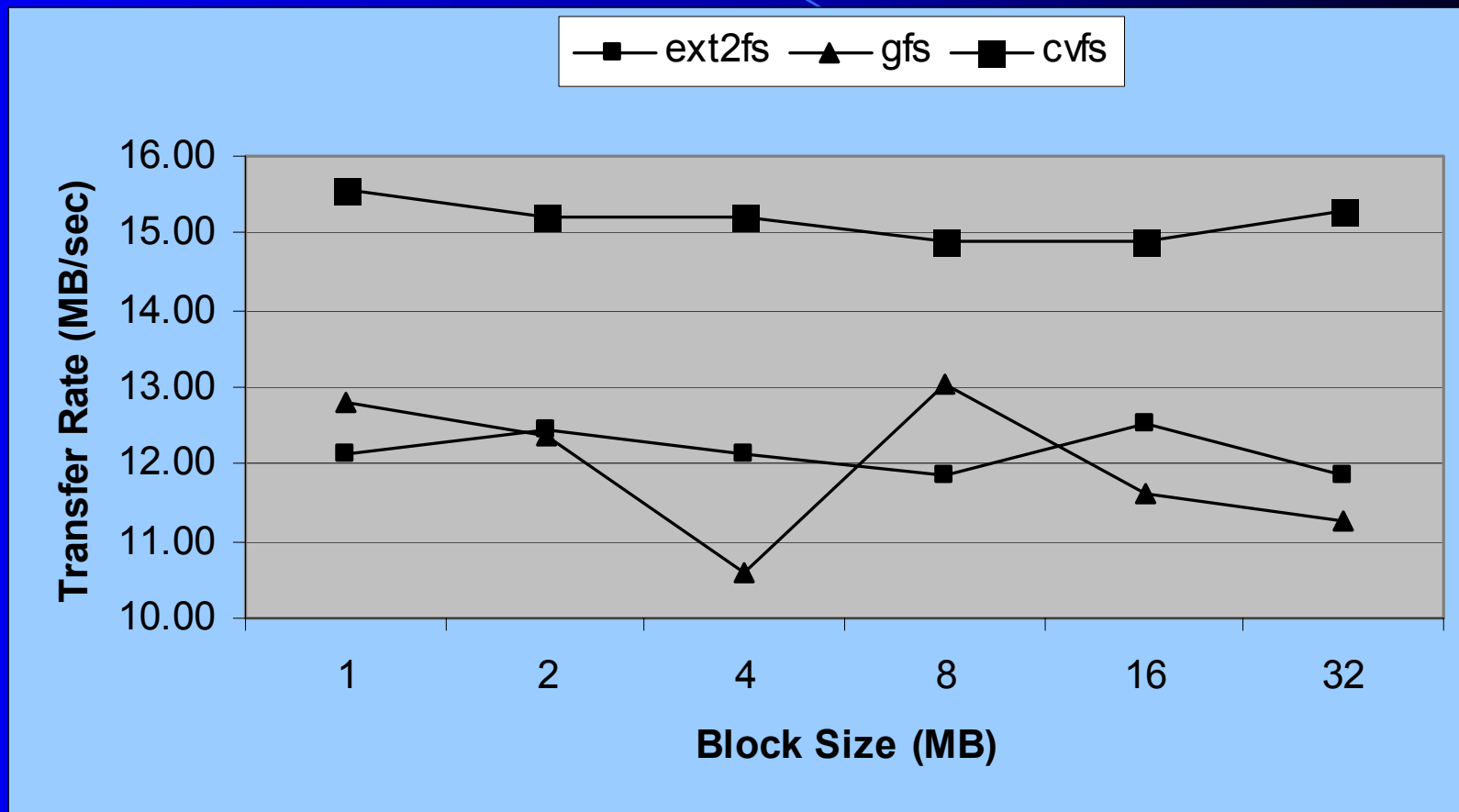
'Linux_G3' Read Performance



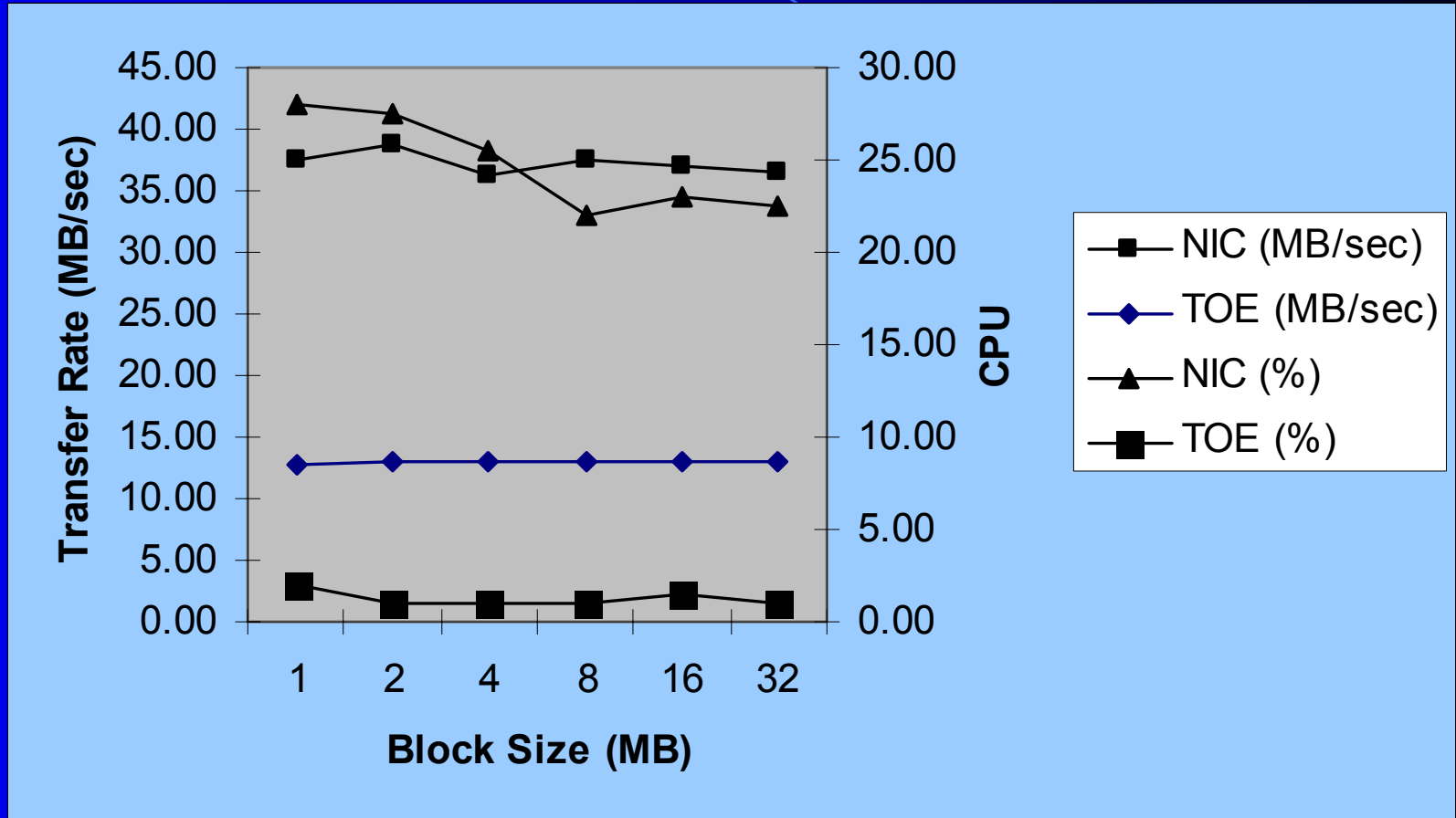
'Linux_MD' Write Performance



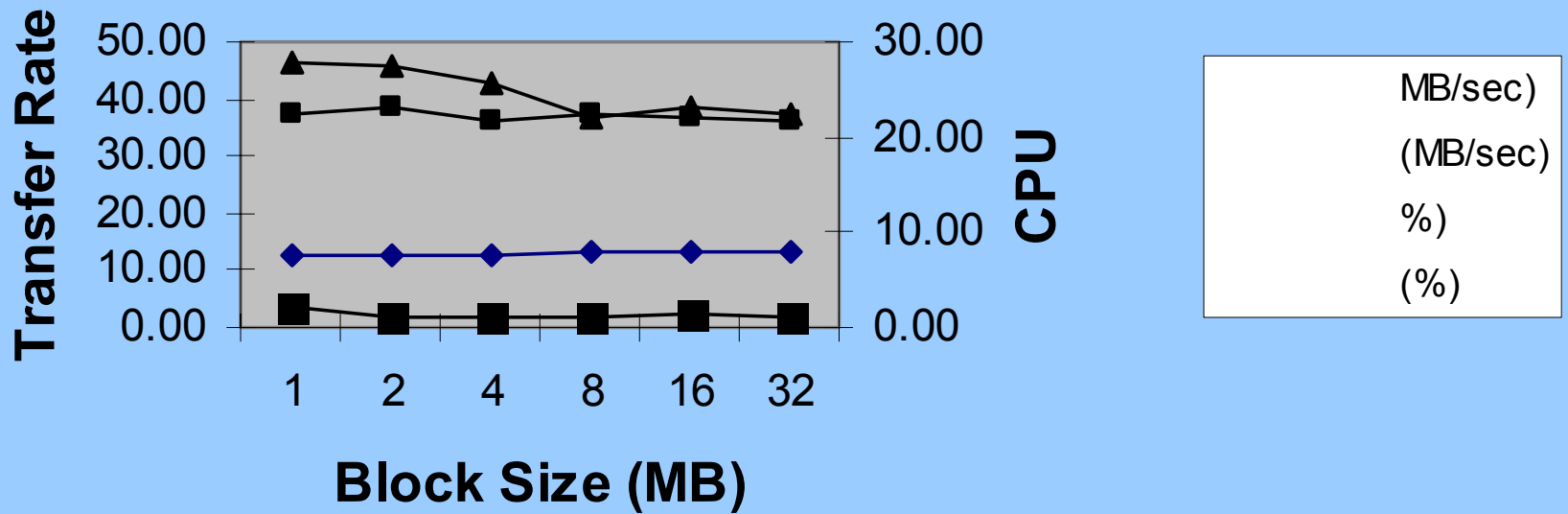
'Linux_MD' Read Performance



TOE Performance - Writes



TOE Performance - Reads



Bonnie++ Results

Number of Files = 10:120:80/44

Chunk Size = 2GB

system	file system	Sequential Output						Sequential Input				Random		Sequential Create						Random Create					
		Per Char		Block		Rewrite		Per Char		Block		Seeks		Create		Read		Delete		Create		Read		Delete	
		KB/sec	% CPU	KB/sec	% CPU	KB/sec	% CPU	KB/sec	% CPU	KB/sec	% CPU	/sec	% CPU	/sec	% CPU	/sec	% CPU	/sec	% CPU	/sec	% CPU	/sec	% CPU	/sec	% CPU
Linux_G2*	ext2fs	9428	99	60735	52	3225	32	9012	98	70047	36	6716.5	45	17420	98	+++++	+++	+++++	+++	17621	99	+++++	+++	+++++	+++
Linux_G2	gfs	8145	97	40740	65	31937	49	8893	98	64626	39	494.6	4	1330	35	+++++	+++	938	55	1230	71	18745	98	952	53
Linux_G2*	cvfs	7584	86	32691	48	4848	32	7647	87	40208	38	260.5	15	24	1	192	17	25	0	24	2	205	18	48	0
Linux_G3	ext2fs	7042	99	41860	42	11340	18	6353	97	16164	12	385.2	4	11231	100	+++++	+++	+++++	+++	12182	99	+++++	+++	+++++	+++
Linux_G3	gfs	5640	97	25467	63	13737	34	6206	95	20683	19	1263.4	19	818	35	15870	100	767	52	437	23	12569	100	731	52
Linux_G3	cvfs	4289	68	10007	28	2114	23	3752	62	7967	19	237.5	22	23	2	146	16	25	0	23	2	98	12	24	1
Linux_MD-IDE	ext2fs	17626	97	45303	31	21270	14	16018	88	44266	13	258.1	1	12623	71	+++++	+++	+++++	+++	15112	90	+++++	+++	8071	28
Linux_MD-iSCSI	ext2fs	11179	58	20761	9	6345	3	10664	65	8678	3	91.1	0	+++++	+++	+++++	+++	15208	13	+++++	+++	+++++	+++	+++++	+++
Linux_MD-iSCSI	gfs	10839	64	23343	20	8680	7	11368	70	13551	5	339.5	2	369	6	+++++	+++	283	4	690	11	+++++	+++	281	4
Linux_MD-iSCSI	cvfs	3277	19	2762	2	753	3	3226	19	7618	7	188.8	7	18	1	82	3	24	0	18	0	74	3	24	0

Postmark Results

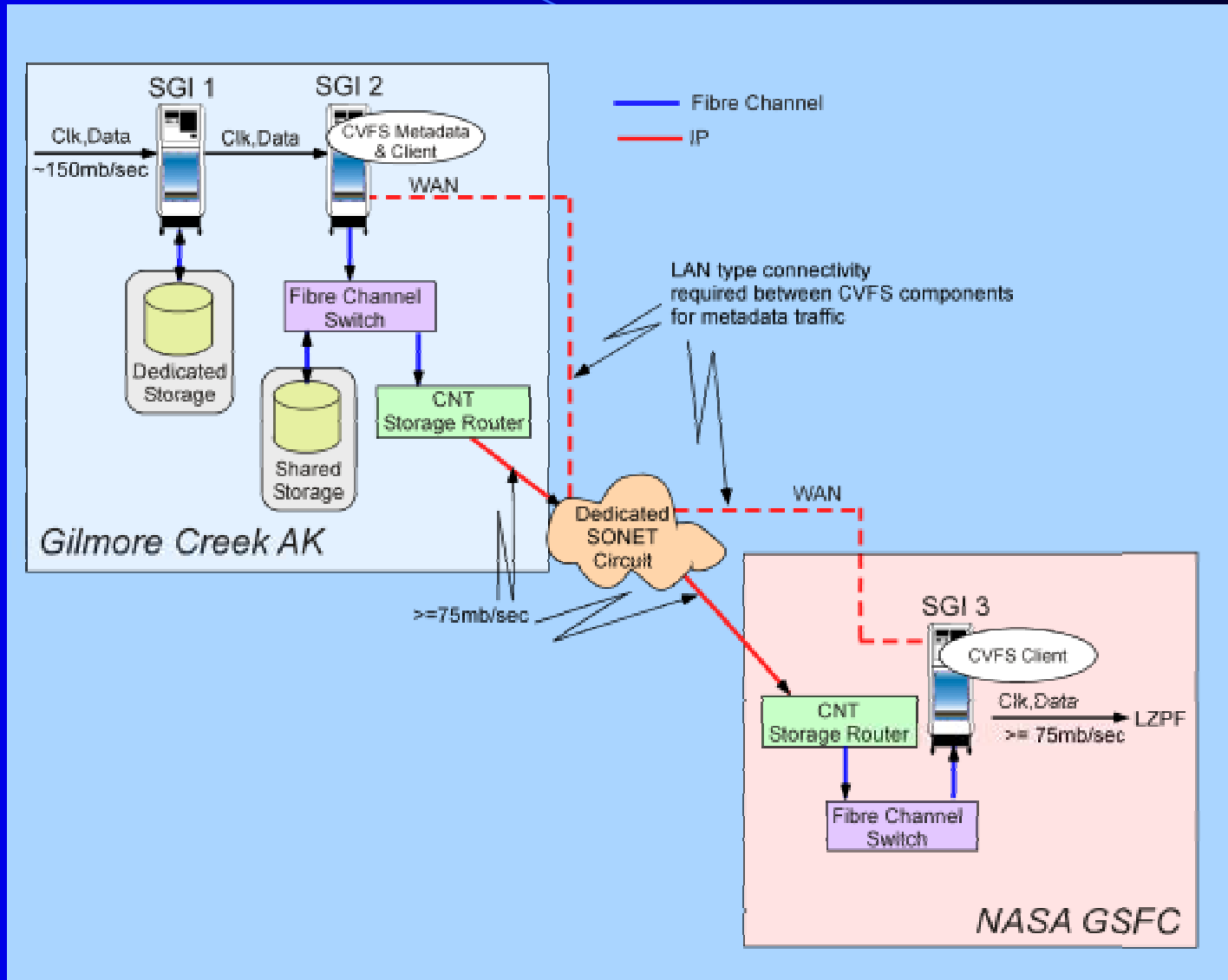
Linux_MD			Time			Files															Data				
File Range	Block Size	FS	Total	Trans	Trans /sec	Crt'd	Crt'd /sec	Crt'd Alone	Crt'd Alone /sec	Mixed w/trans	Mixed w/trans /sec	Read	Read /sec	App	App /sec	Del	Del /sec	Del Alone	Del Alone /sec	Mixed w/trans	Mixed w/trans /sec	MB read	MB Read /sec	MB Written	MB Wrirtten /sec
16K - 1M	512	ext2	3	1	500	730	243	500	250	230	230	252	252	248	248	730	243	460	460	270	270	139.76	46.59	423.93	141.31
16K - 1M	512	gfs	42	15	33	730	17	500	83	230	15	252	16	248	16	730	17	460	21	270	18	139.76	3.33	423.93	10.09
16K - 1M	512	cvfs	204	92	5	730	3	500	5	230	2	252	2	248	2	730	3	460	27	270	2	139.76	702(K)	423.93	2.08
1M - 8M	4096	ext2	238	140	3	742	3	500	5	242	1	252	1	248	1	742	3	484	484	258	1	1210.73	5.09	3604.88	15.15
1M - 8M	4096	gfs	283	154	3	742	2	500	4	242	1	252	1	248	1	742	2	484	26	258	1	1210.73	4.28	3604.88	12.74
1M - 8M	4096	cvfs	1126	505	0	500	0	500	0	242	0	252	0	248	0	742	0	484	25	258	0	1210.73	1.08	3604.88	3.2

Linux_G2			Time			Files															Data				
File Range	Block Size	FS	Total	Trans	Trans /sec	Crt'd	Crt'd /sec	Crt'd Alone	Crt'd Alone /sec	Mixed w/trans	Mixed w/trans /sec	Read	Read /sec	App	App /sec	Del	Del /sec	Del Alone	Del Alone /sec	Mixed w/trans	Mixed w/trans /sec	MB read	MB Read /sec	MB Written	MB Wrirtten /sec
16K - 1M	512	ext2	6	2	250	730	121	500	166	230	115	252	126	248	124	730	121	460	460	270	135	139.76	23.29	423.93	70.66
16K - 1M	512	gfs	19	8	62	730	38	500	83	230	28	252	31	248	31	730	38	460	92	270	33	139.76	7.36	423.93	22.31
16K - 1M	512	cvfs	90	37	13	730	8	500	13	230	6	252	6	248	6	730	8	460	27	270	7	139.76	1.55	423.93	4.71
1M - 8M	4096	ext2	93	51	9	742	7	500	12	242	4	252	4	248	4	742	7	484	484	258	5	1210.73	13.02	3604.88	38.76
1M - 8M	4096	gfs	147	75	6	742	5	500	7	242	3	252	3	248	3	742	5	484	121	258	3	1210.73	8.24	3604.88	24.52
1M - 8M	4096	cvfs	318	176	2	742	2	500	4	242	1	252	1	248	1	742	2	484	25	258	1	1210.73	3.81	3604.88	11.34

Application Testing – U of MD

- Successfully recognized and mounted storage located ~6 miles away
- Generated MODIS composite data
 - ext2fs > 1hr 45 min
 - Plus ftp overhead time of 45 min
 - gfs > 2 hr 8 min
 - cvfs > 3hr 15 min

AK-To-MD FCIP Test



Conclusions

- Technology is straightforward to implement and test
- IP is more than just a viable SAN technology
 - Poised to have a dramatic impact
 - Market has yet to play out the options - iSCSI, FCIP and/or iFCP
 - Vendor commitment still forming
- Gain in flexibility offsets bandwidth loss for potentially a large category of users

