Two Case Studies of the Application of Dynamic Modeling Techniques in Performance Assessment and Prediction of Complex Shared Storage Architectures

Marti Bancroft, Phillip L. (Rocky) Snyder, and Mark Woodyard

Sun Microsystems, Palo Alto, California, USA

HPC Technology Engineering

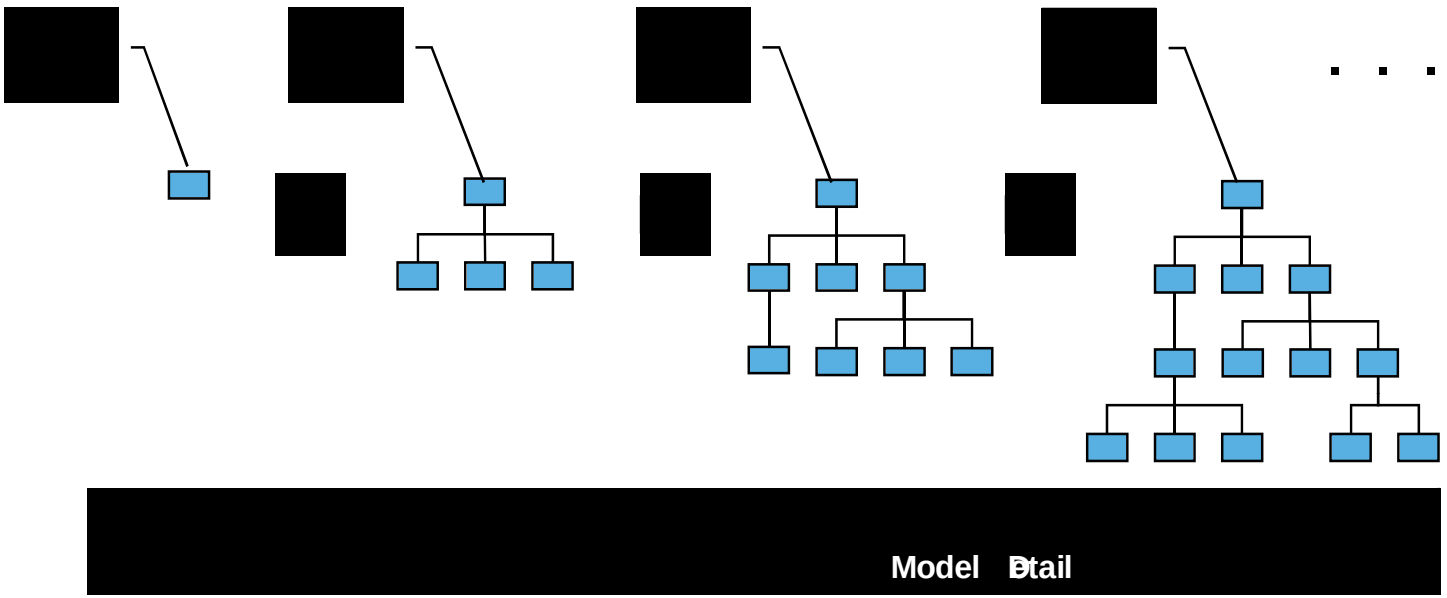
HPC Horizontal Solutions

Outline

- Role of dynamic models in predicting the behavior of large systems
- Systems being studied
- Test data which invalidated our models
- Next steps

Overview of Findings

- Oops! (no, not Object Oriented Programming System, but… new data invalidated our models - an oops!)
- Customer calls this “frictionless I/O”
- Key question: how far from ‘local’ will this be valid?
- Key question: at what load will this return to “normal”? I.e. serious reduction in efficiency...



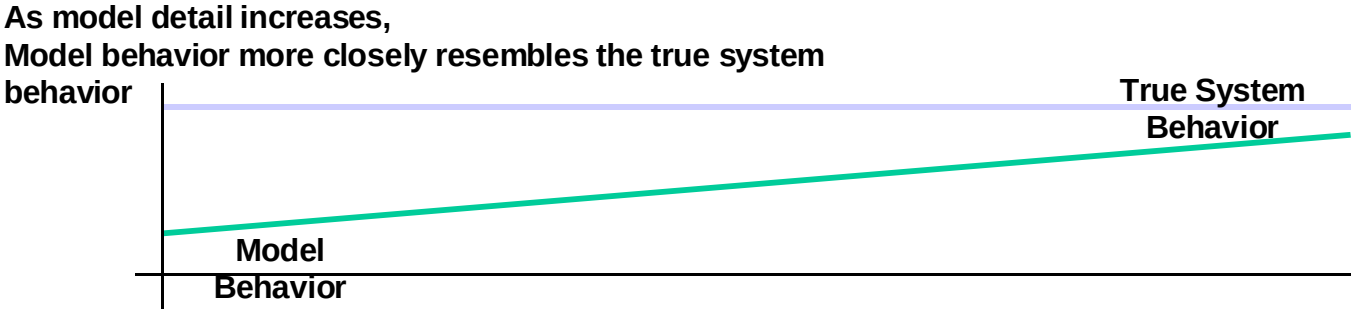


Figure 3 Evolution of a Dynamic Model Over Time

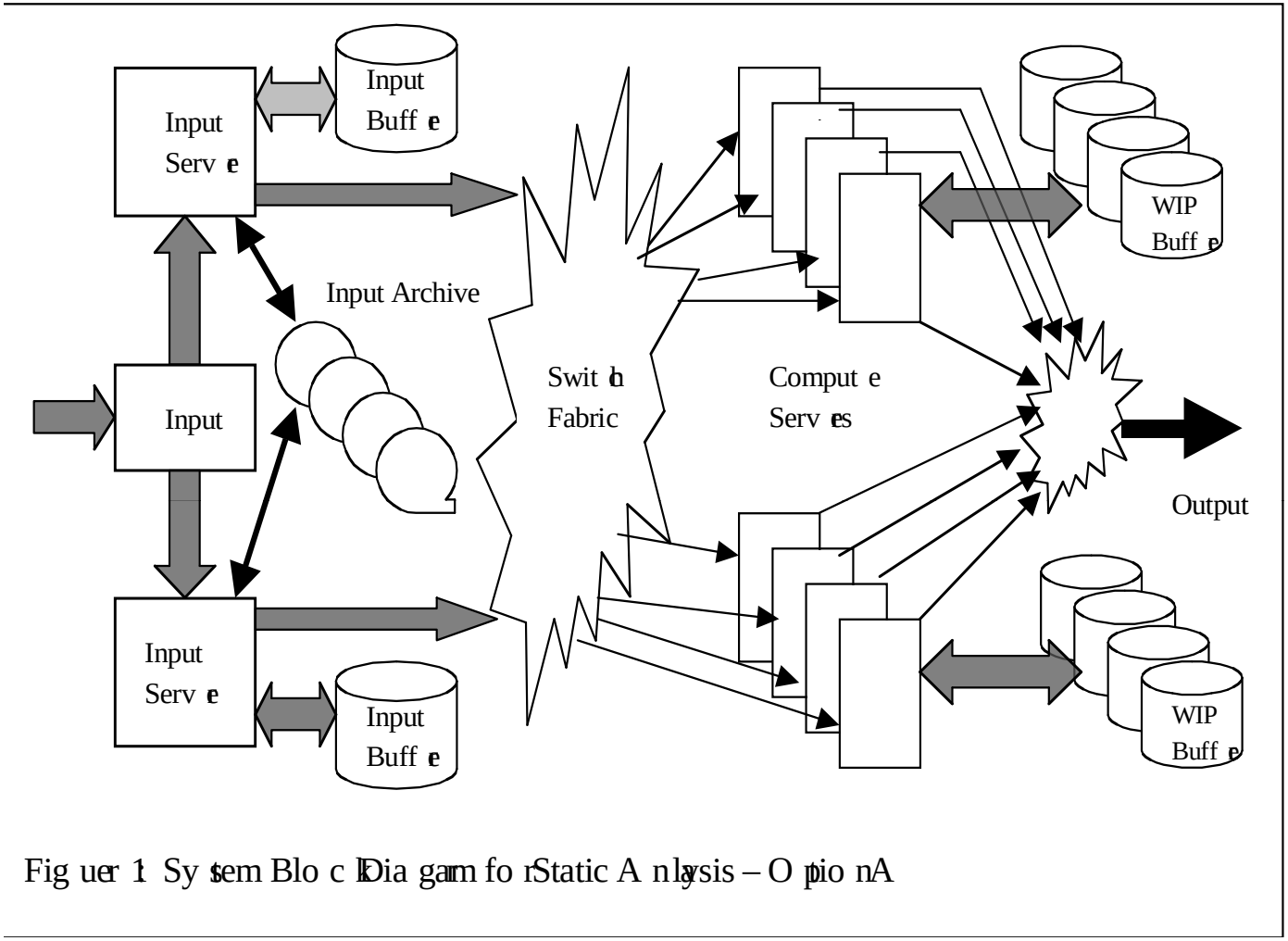


Figure 1 System Block Diagram for Static Analysis – Option A

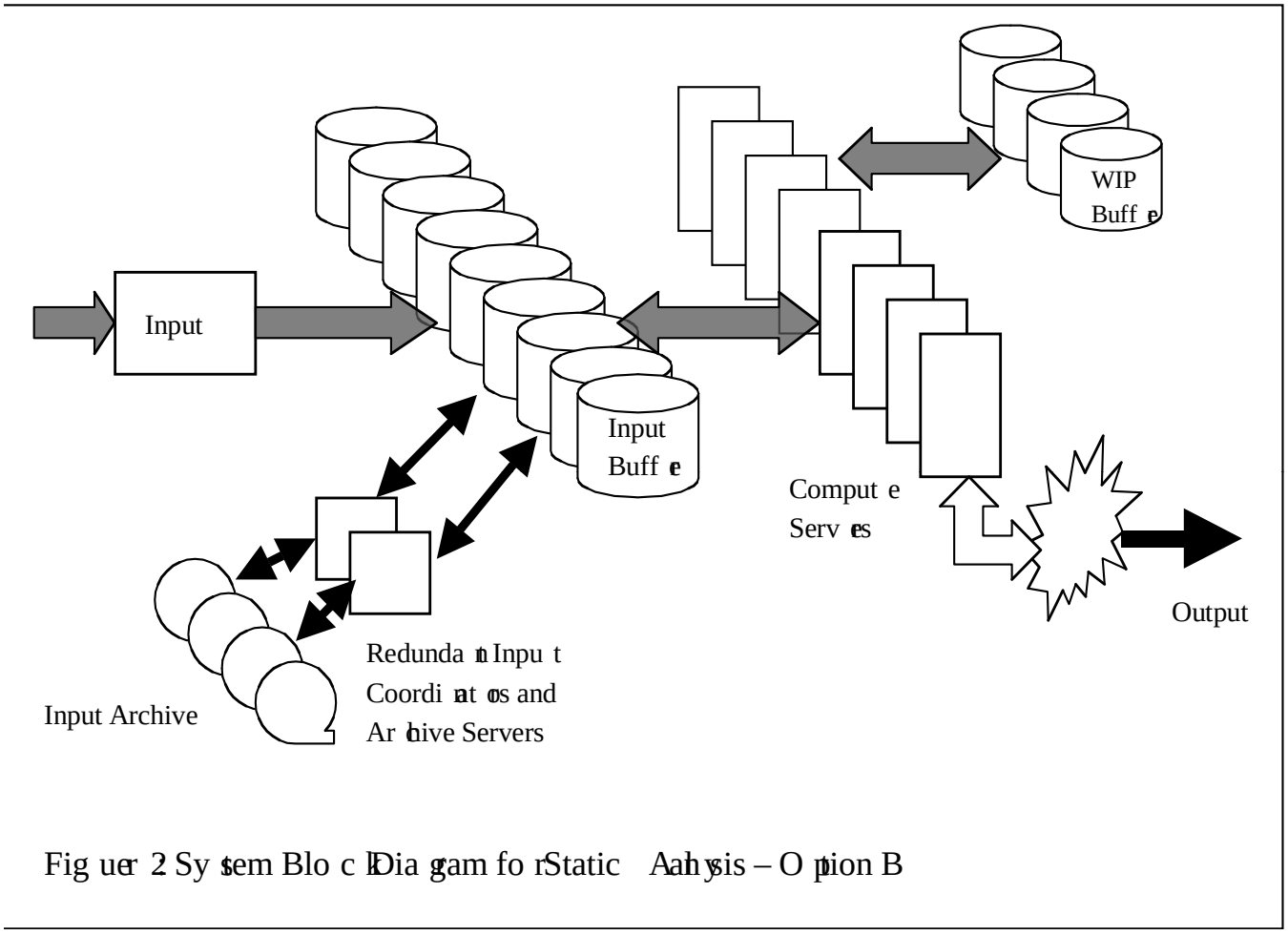


Figure 2 System Block Diagram for Static Analysis – Option B

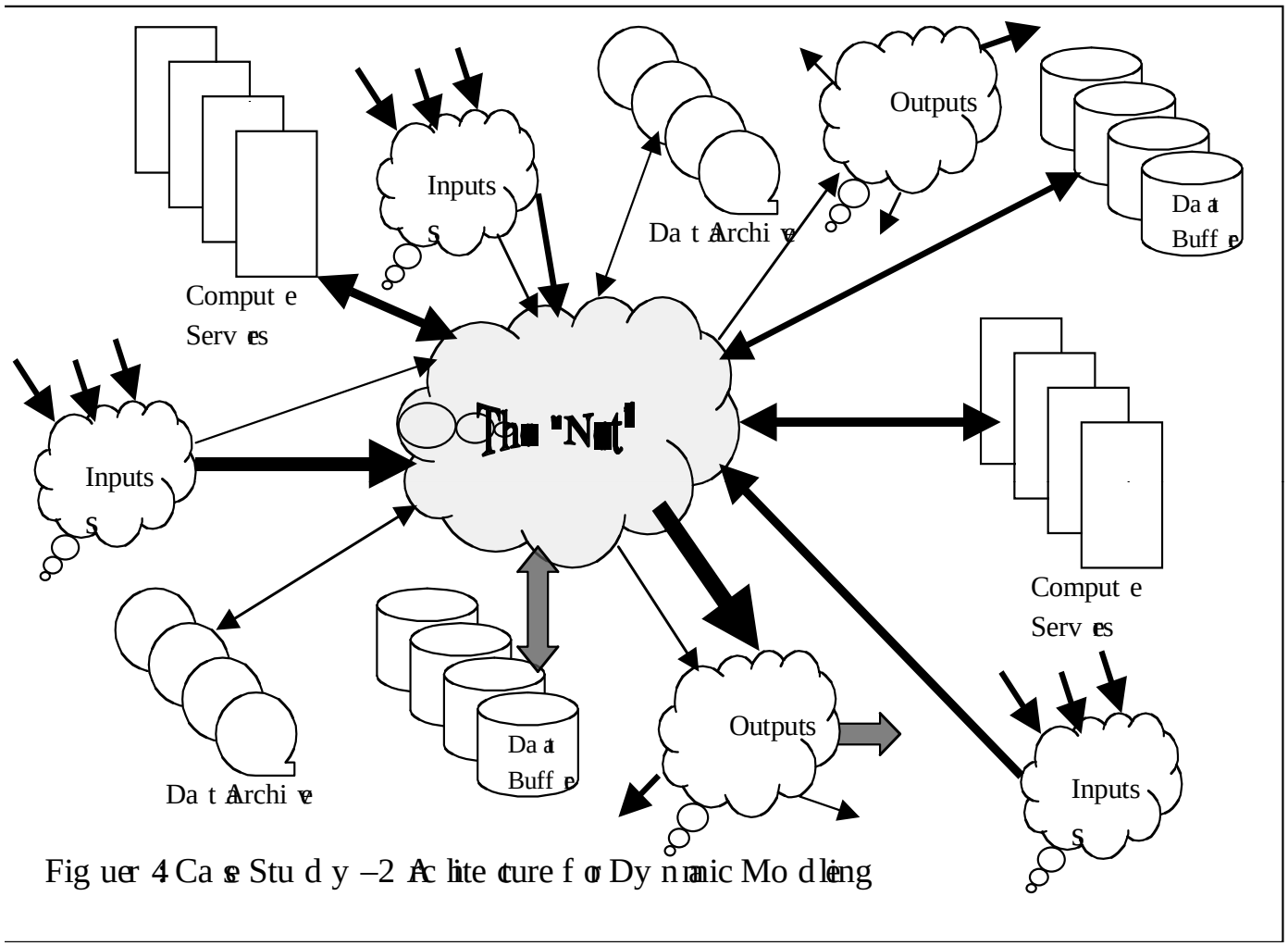
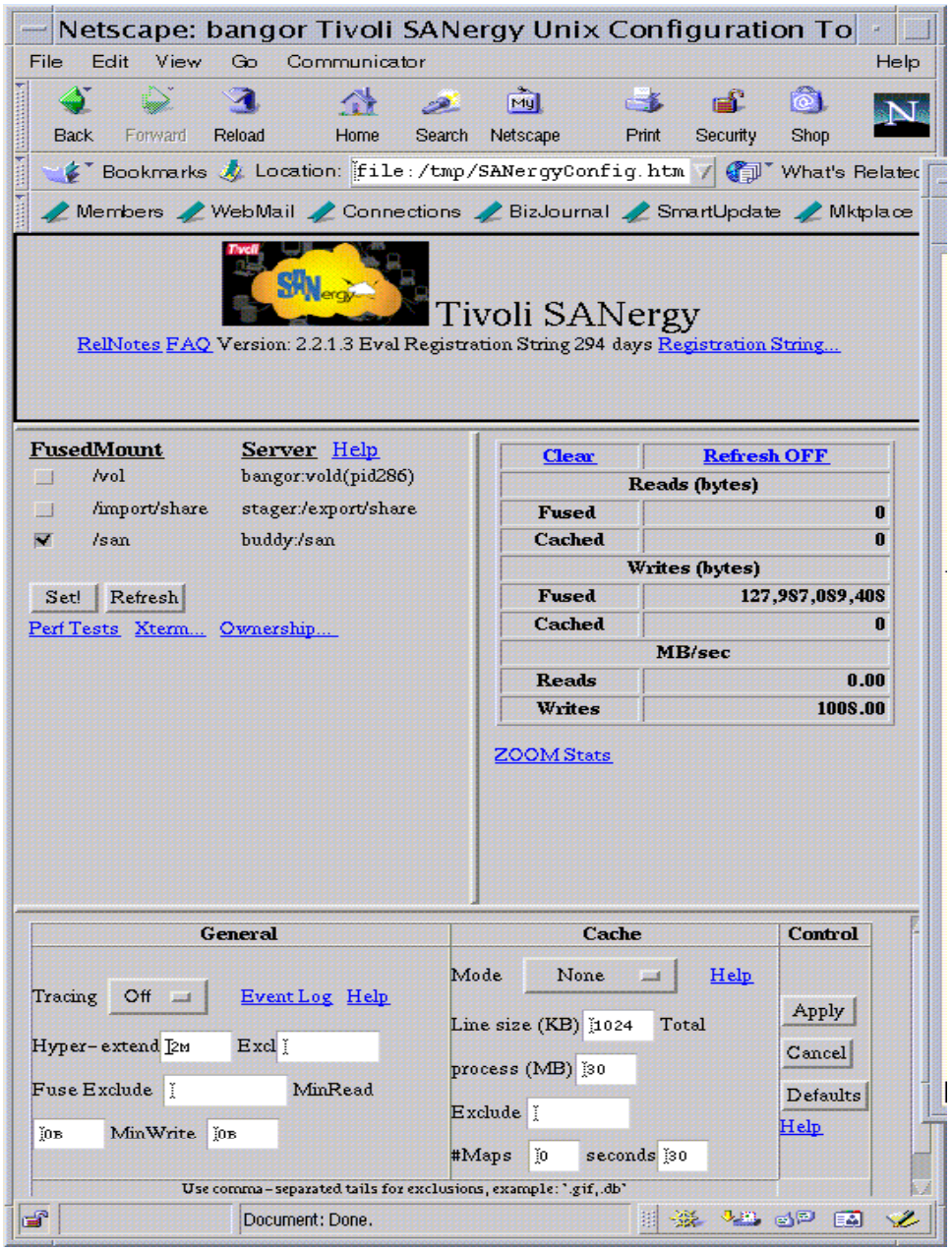


Figure 4 Case Study – 2 Architecture for Dynamic Modeling

SANergy client writes to 105GB file, preallocated, sustaining more than 1 Gigabyte per second, using 12 threads (LSC's recommended minimum) (1008 MB/s shown)

(QFS 3.5.0-28b)

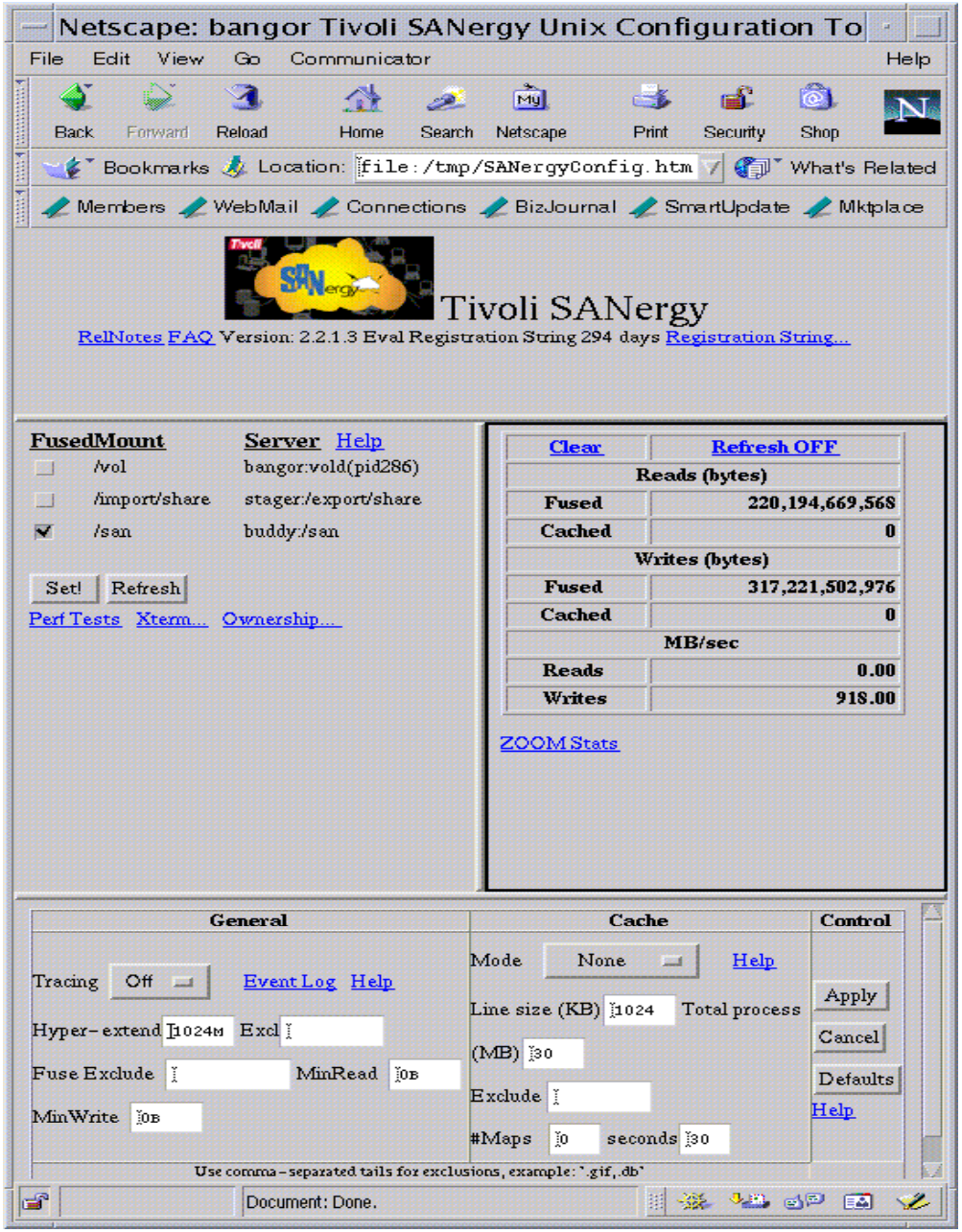


SANergy client writes to 105GB file, 18 cpus, preallocated, sustaining more than 1 Gigabyte per second, using 12 threads (LSC's recommended minimum) (1008 MB/s on screen)

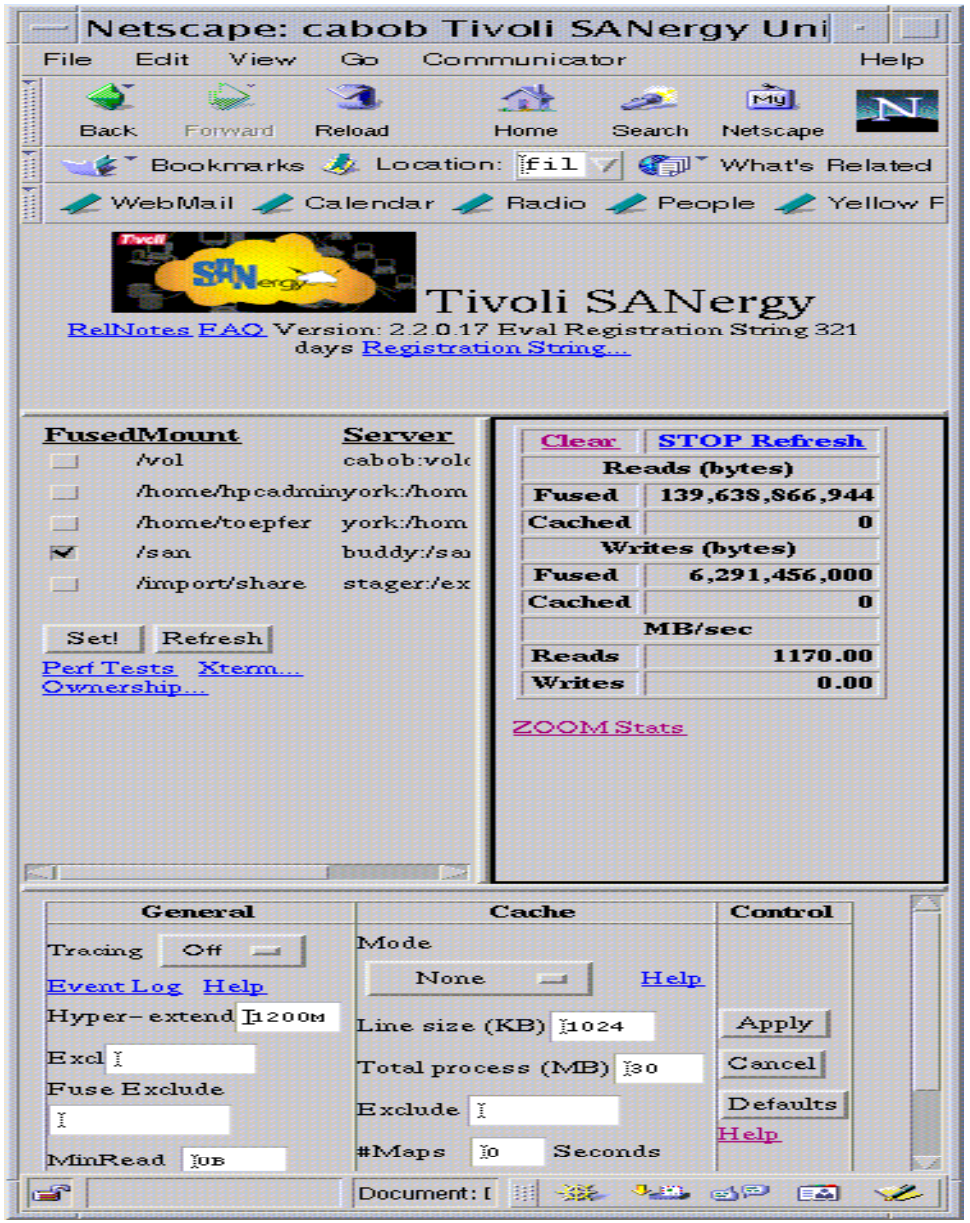
IOSTAT:

	extended device statistics									tty		cpu			
device	r/s	w/s	kr/s	kw/s	wait	actv	svc_t	%w	%b	tin	tout	us	sy	wt	id
sd0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0	0	652	5	3	91	0
ssd0	0.0	161.6	0.0	82738.4	0.0	10.0	61.8	0	100						
ssd1	0.0	161.6	0.0	82738.4	0.0	9.1	56.5	0	100						
ssd2	0.0	161.2	0.0	82533.6	0.0	9.4	58.1	0	100						
ssd3	0.0	161.8	0.0	82840.8	0.0	9.3	57.6	0	100						
ssd4	0.0	162.0	0.0	82943.2	0.0	9.3	57.6	0	100						
ssd5	0.0	161.8	0.0	82840.8	0.0	8.4	52.1	0	100						
ssd6	0.0	161.8	0.0	82840.8	0.0	9.0	55.5	0	100						
ssd7	0.0	162.2	0.0	83045.6	0.0	8.8	54.6	0	100						
ssd8	0.0	161.8	0.0	82840.8	0.0	8.7	53.5	0	100						
ssd9	0.0	161.4	0.0	82636.1	0.0	9.4	58.4	0	100						
ssd10	0.0	161.6	0.0	82738.5	0.0	8.9	55.2	0	100						
ssd11	0.0	162.0	0.0	82943.3	0.0	8.4	51.6	0	100						

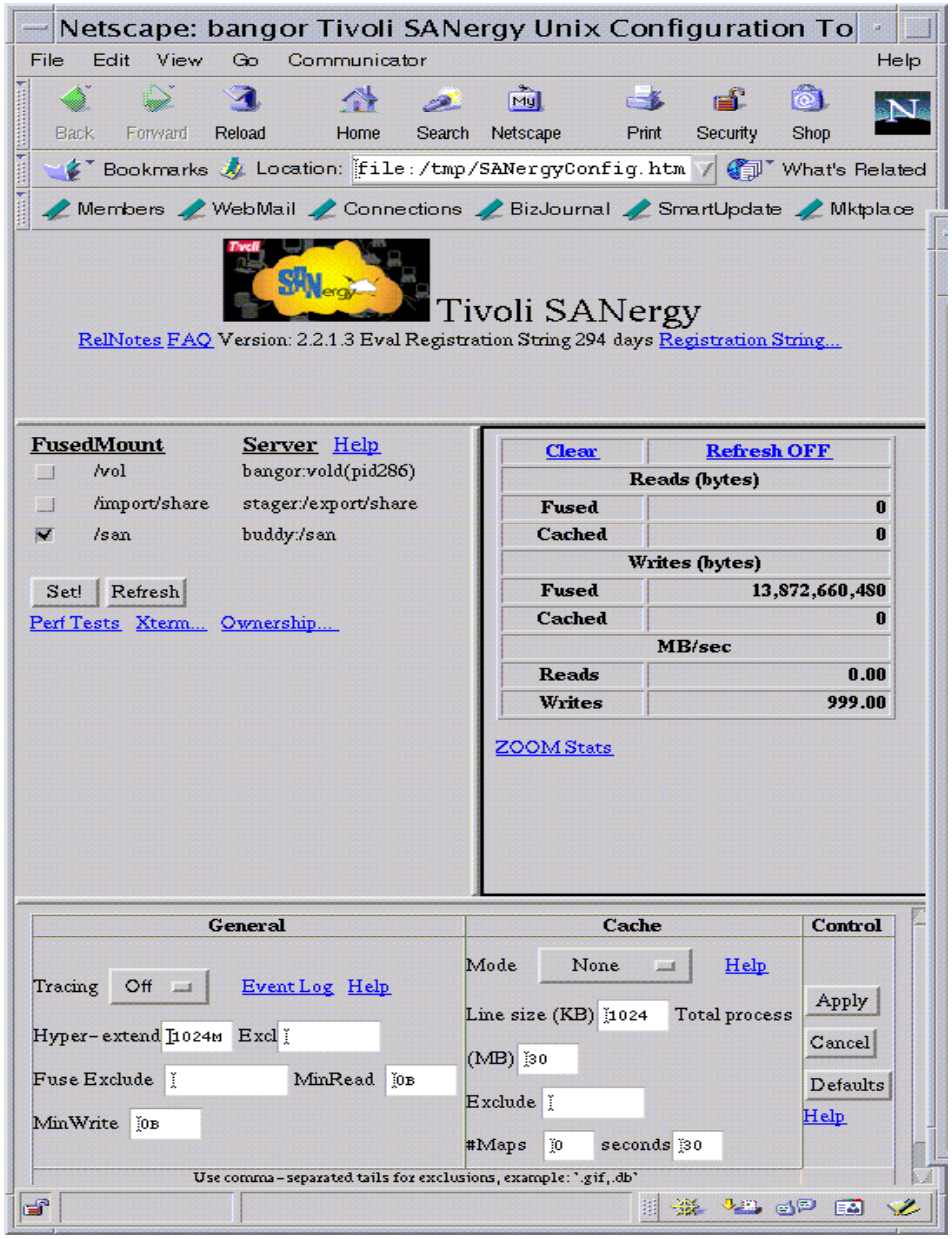
SANergy client writes to 105GB file, preallocated, 6 threads (0.5 per T3 brick) (918 MB/s shown)



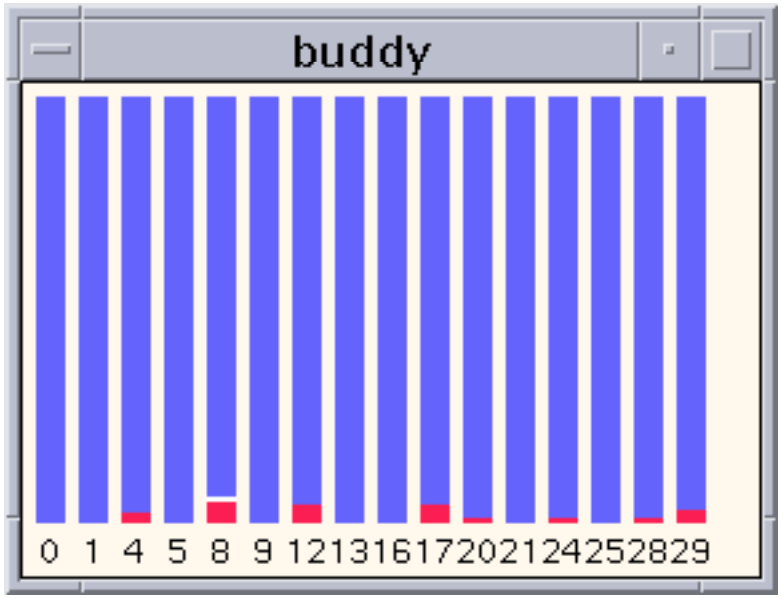
SANergy client reads from 105GB file, 6 threads (0.5 per T3 brick) - using prerelease version at customer demos - shown bursting at 1170 MB/s



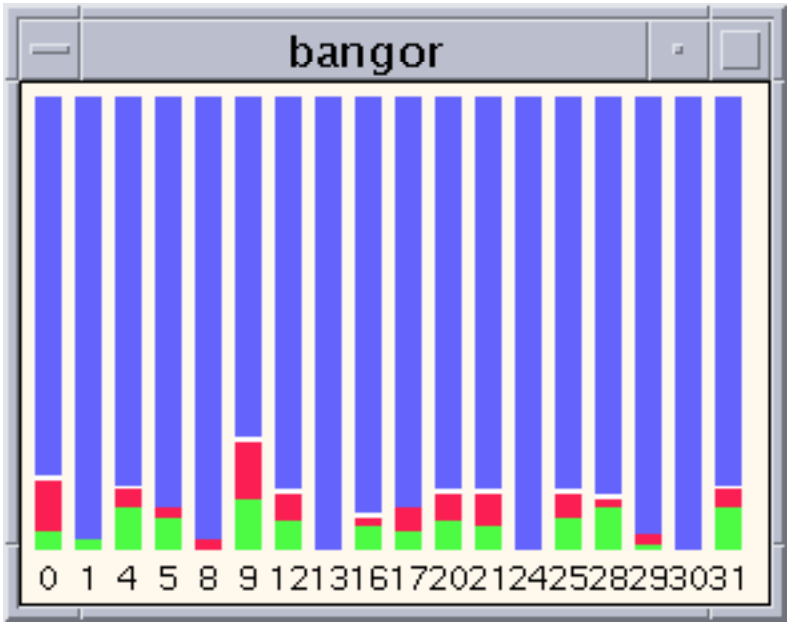
SANergy client writes to 105GB file, 12 threads but without file preallocation, 1GB extents, showing almost 1 Gigabyte per second (999 MB/second)



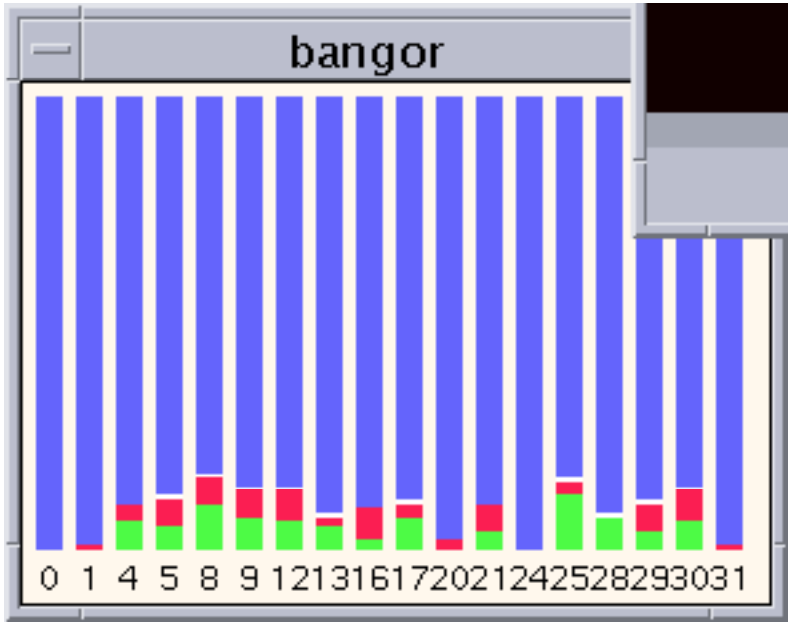
SANergy client writes to 105GB file, 12 threads but without file preallocation, 1GB extents, CPU load on metadata controller during 999 MB/s stream



SANergy client writes to 105GB file, 12 threads but without file preallocation, 1GB extents, CPU load on client during 999 MB/s stream



SANergy client writes to 105GB file, 12 threads with file preallocation, CPU load on client during > 1 GB/second stream



SANergy 2.2.0.17 CLIENT AND QFS 3.5.0.22 NATIVE COMBINATION TESTING USING VERY LARGE FILES

Host Name	Test Used	Block/ File Size	Threads	Sustained	Aggregate
SANergy client reads on caboa-baseline	jdvtto/san on 12 T3s (1 file per process)	6MB/24 GB	R=12 (q=24)	R - 1106.64 - 1108.66 MB/s	N/A (tests alone)
		6MB/ 105GB	R=12 (q=24) R=32 (q=64)	R - 1091.06 - 1104.90 MB/S R - 1112.18 MB/s	N/A (tests alone)
		3MB/ 105GB	R=12 (q=24)	R - 1072.64 - 1106.29 MB/s	N/A (tests alone)
SANergy client reads on cabob-baseline	same	6MB/ 105GB	R=32 (q=64)	R - 1103.96 MB/s	N/A (tests alone)
SANergy client reads on caboa and cabob	jdvtto/san on 12 T3s (1 file per process)	6MB/ 105GB	R=32 (q=64)	R=53997 + 695.30 R=492.11 + 617.34	123527 MB/s 110945 MB/s
SANergy client reads on caboa and cabob; QFS native writes on buddy	jdvtto/san on 12 T3s (1 file per process)	6MB/ 105GB	W=32 (q=64) R=12 (q=24)	W=408.56 MB/s R=392.46 MB/s R=468.32 MB/s	126934 MB/s
SANergy client reads -baseline	dvtto /san on 12 T3s	6MB/ 105GB	R=6 (q=12)	caboa: R=1108.22 MB/s; cabob: R=1100.33 MB/s	N/A (tests alone)
QFS baseline writes	dvtto /san on 12 T3s	6MB/ 105GB	W=12 (q=24)	buddy: W=798.89, 795.93, 793.27 MB/s	N/A (tests alone)
SANergy client reads on caboa and cabob	dvtto /san on 12 T3s	6MB/ 105GB	R=6 (q=12)	R=52055 + 520.66 MB/s	104121 MB/s

SANergy 2.2.0.17 CLIENT AND QFS 3.5.0.22 NATIVE COMBINATION TESTING
USING VERY LARGE FILES

Host Name	Test Used	Block/ File Size	Threads	Sustained	Aggregate
SANergy client reads on c drive and cache b; QFS native writes on buddy	dvt to /san on 12 T3s	6MB/ 105GB	W= 12 (q =24) R =6 (q =12)	W= 450.48 MB/s R +331 11 + 332. 16 MB/s	1113 75 MB/s
QFS baseline read from one file system, write to second (1 process)	jdvt from /capture (6 x T3s) to /san (12 x T 3s)	6MB/ 105GB	R =32, W= 32 (q =64)	Sustained= free civ e write rate to /san= 518. 21 - 526.25 MB/s	N/A (tests alone)
SANergy client reads on c drive and cache b	dvt to /san on 12 T3s	6MB/ 105GB	R =6 (q =12) (4 processes, 4 files)	R =216 34 + 217. 93 + 216. 41 + 216.88 MB/s	867. 56 MB/s
SANergy client reads on c drive and cache b; QFS native writes on buddy	dvt to /san (12x T3s) = writes; 2 reads from each client	6MB/ 105GB	W= 12 (q =24) R =6 (q =12)	W= 279.59 MB/s R +184 48 + 183. 22 + 182. 94 + 183.32	1013 55 MB/s

METADATA LAN LOADING EFFECT OF RATES

(SANergy 22.0.17 and QFS 3.5.0.22)

Host Name	Test Used	Block/ File Size	Threads	Sustained	Aggregate
SANergy client reads on caboa - baseline - unloaded metadata path	dvt to /san on 12 T3s (1 file	6MB/ 105GB	R=6 (q=12)	R - 1109.29 MB/s	N/A (tests alone)
SANergy client reads on cabob - baseline - unloaded metadata path	same	6MB/ 105GB	R=6 (q=12)	R - 1110.30 MB/s	N/A (tests alone)
QFS baseline read fom one file system, write to second (1 process)	jdvt from /capture (6 x T3s) to /san (12 x T3s)	6MB/ 105GB	R=32, W=32 (q=64)	Sustained= effective write rate to /san= 518.21 - 531.12 MB/s	N/A (tests alone)
SANergy client reads on caboa and cabob; QFS native writes on buddy from /capture to /san - baseline - unloaded metadata path	dvt from /san on 12 T3s (2 files); jdvt from /capture (6 x T3s) to /san (12 x T3s)	6MB/ 105GB	R=32, W=32 (q=64) R=6 (q=12)	W +501.60 MB/s R = 330.88 + 331.47 MB/s	1163.95 MB/s
SANergy client read on caboa - heavy load on metadata path	dvt to /san on 12 T3s	6MB/ 105GB	R=6 (q=12)	R = 785.24 MB/s	N/A (tests alone)
SANergy client reads on caboa and cabob; QFS native writes on buddy from /capture to /san - baseline - heavy load on metadata path	dvt from /san on 12 T3s (2 files); jdvt from /capture (6 x T3s) to /san (12 x T3s)	6MB/ 105GB	R=32, W=32 (q=64) R=6 (q=12)	W =491.50 MB/s sustained R = 306.63 + 291.65 MB/s	1089.51 MB/s