

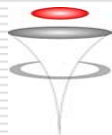
SAS SPDS

Performance Tuning in UNIX

Twin Cities Area SAS User Group Meeting
Dec 11, 2008

Basha Shaik

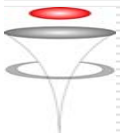
Basha.Shaik@gmail.com



IntelOpt LLC

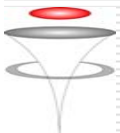
Performance Tuning

- ❑ Performance is user expectations of a job completion in a given system
- ❑ Performance tuning is tweaking the system parameters (Hardware, OS, Application and the job) to achieve the user's expectations
- ❑ It is an iterative process



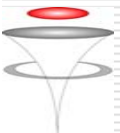
Performance Tuning Situation

- ❑ Base job is faster than SPDS
- ❑ Unpredictable runtime
- ❑ Variance in real CPU% and User+System CPU%



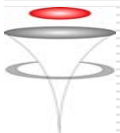
What do we check

- CPU
- Memory
- IO
- Disk
- Application
- Databases (Queries, Connections)



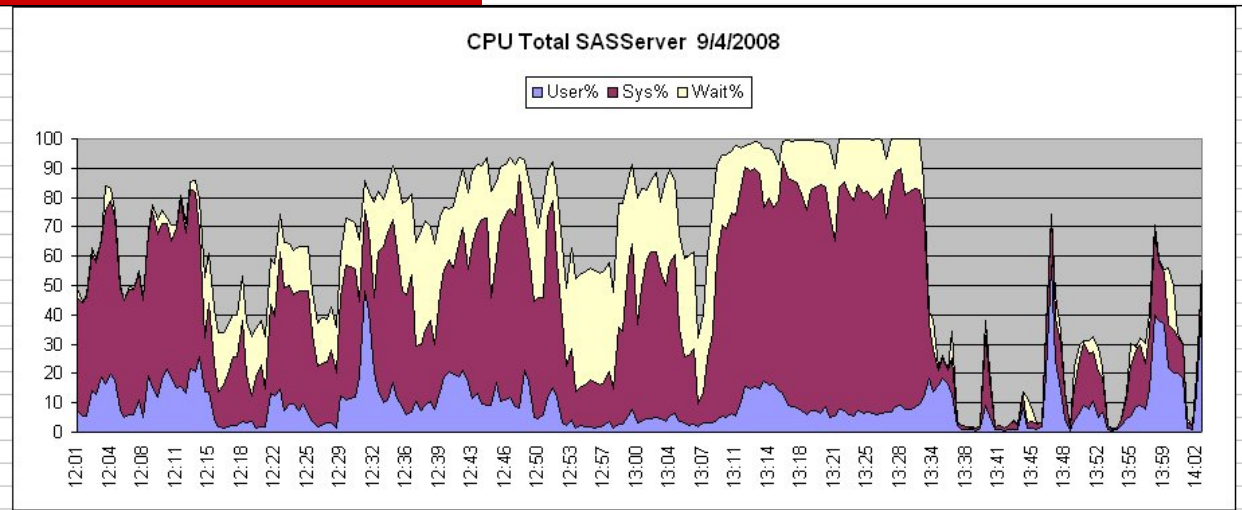
Performance Monitors

- ❑ What do we use to monitor the CPU, Memory, IO, Disk utilization etc.
 - vmstat
 - netstat
 - iostat
 - topas
 - nmon
 - ❑ Usage: `nmon -f -s 30 -c 300 -t`
 - ❑ Nmon analyser:
http://www.ibm.com/developerworks/aix/library/au-nmon_analyser/
 - Sar
 - -dtrace (on solaris)
 - Options fullstimer in SAS Code

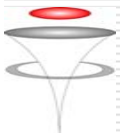
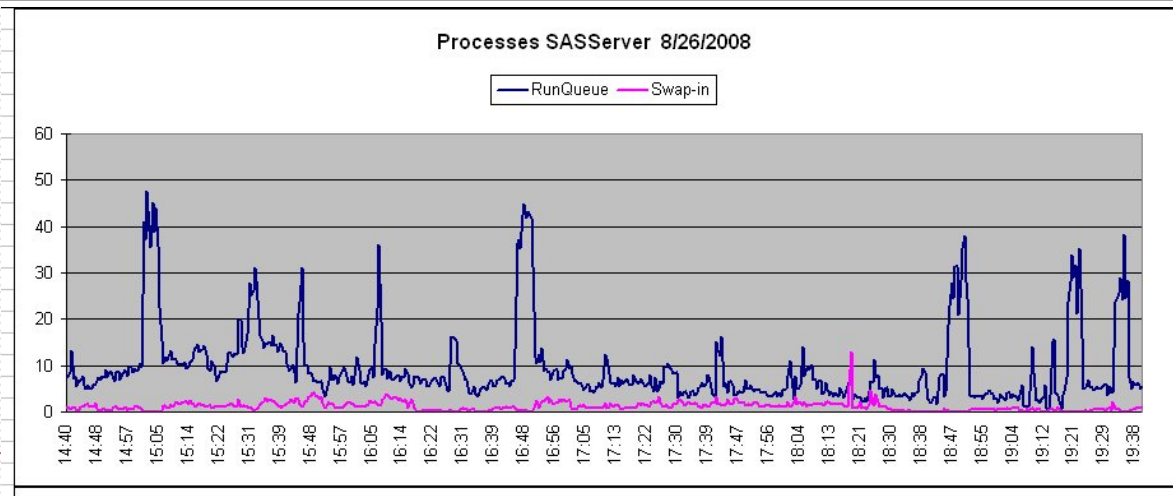


CPU Usage

- User %
- System %
- Wait%



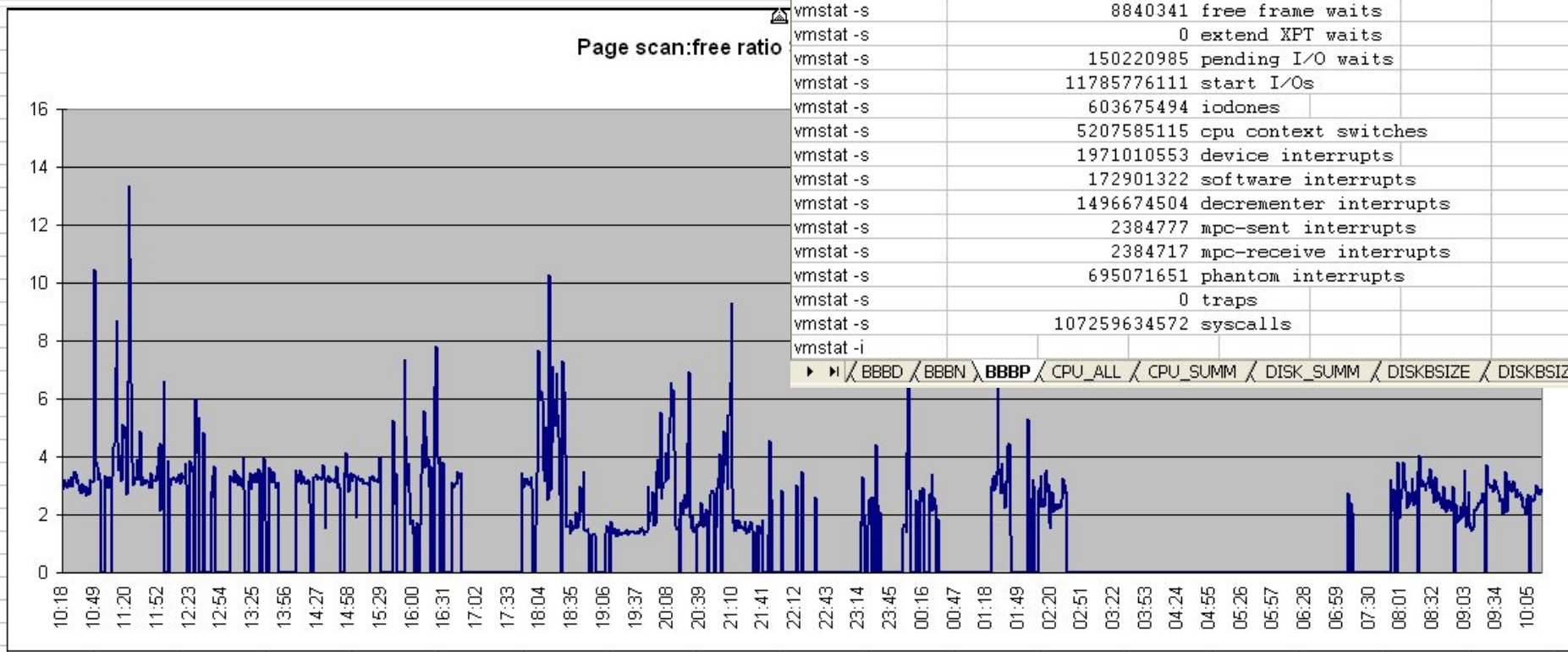
- Runqueue Depth



Memory

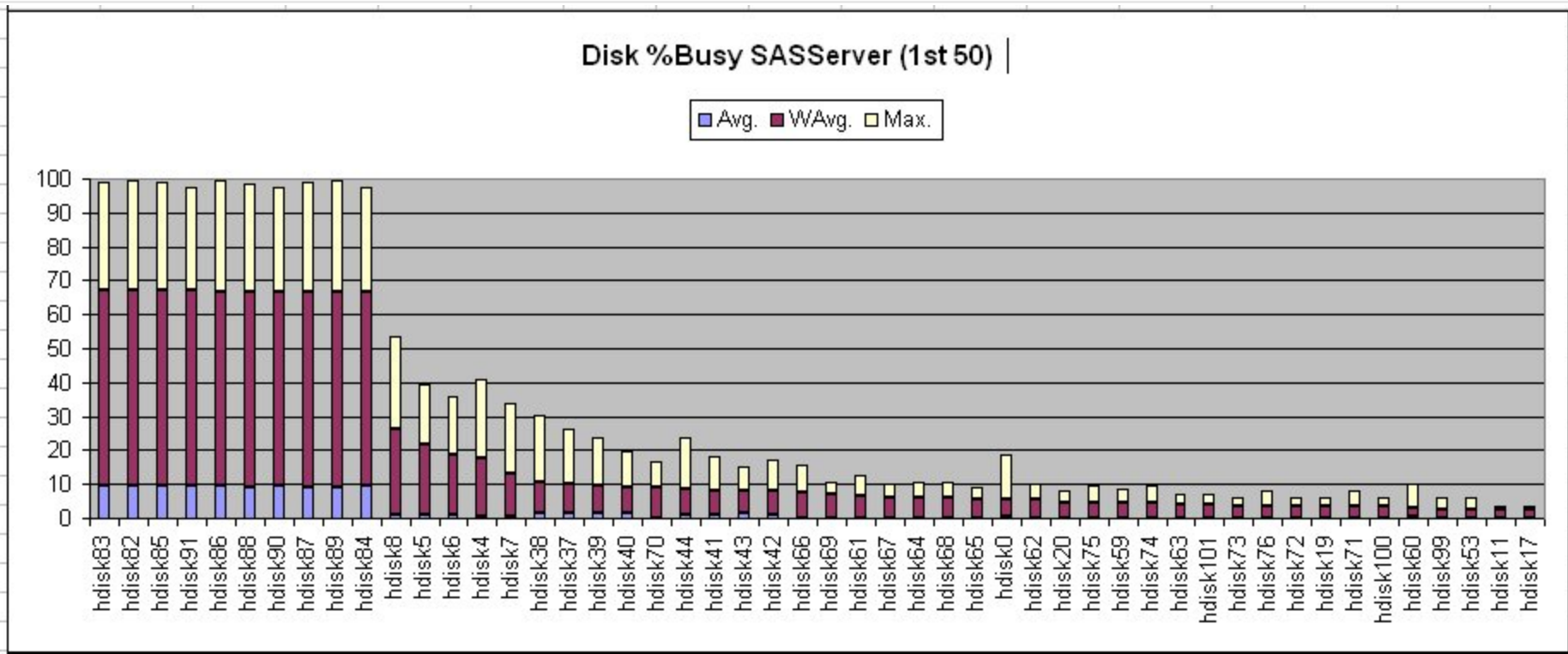
- ❑ Page Scan Rate (vmstat)
- ❑ Swapping Activity (iostat)

vmstat -s	3781375411	total address trans. faults		
vmstat -s	7184512054	page ins		
vmstat -s	4644880093	page outs		
vmstat -s	203009	paging space page ins		
vmstat -s	1153696	paging space page outs		
vmstat -s	0	total reclaims		
vmstat -s	1650301143	zero filled pages faults		
vmstat -s	20572	executable filled pages faults		
vmstat -s	28402938186	pages examined by clock		
vmstat -s	19718	revolutions of the clock hand		
vmstat -s	9473171833	pages freed by the clock		
vmstat -s	26908251	backtracks		
vmstat -s	8840341	free frame waits		
vmstat -s	0	extend XPT waits		
vmstat -s	150220985	pending I/O waits		
vmstat -s	11785776111	start I/Os		
vmstat -s	603675494	iodones		
vmstat -s	5207585115	cpu context switches		
vmstat -s	1971010553	device interrupts		
vmstat -s	172901322	software interrupts		
vmstat -s	1496674504	decrementer interrupts		
vmstat -s	2384777	mpc-sent interrupts		
vmstat -s	2384717	mpc-receive interrupts		
vmstat -s	695071651	phantom interrupts		
vmstat -s	0	traps		
vmstat -s	107259634572	syscalls		
vmstat -i				
▶ / BBBB / BBBN \ BBBP / CPU_ALL / CPU_SUMM / DISK_SUMM / DISKBSIZE / DISKBSIZE1				



IO and Disk

- Check throughput
- Disk %busy



Monitoring tool - topas

```

Topas Monitor for host:      apsa0019
Thu Dec 11 10:34:52 2008   Interval:  2

Kernel  64.7  |#####|
User    34.9  |#####|
Wait    0.3  |#      |
Idle    0.1  |#      |
Physc = 2.81                $Entc= 140.5

Network  KBPS  I-Pack  O-Pack  KB-In  KB-Out
lo0      361.8  26.5    26.5    180.9  180.9
en4      1.9    4.5     4.5     0.4    1.5
en5      0.0    0.0     0.0     0.0    0.0

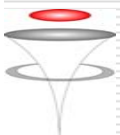
Disk     Busy%  KBPS    TPS  KB-Read  KB-Writ
hdisk4   96.6  106.8K  1.3K  2.4K    104.4K
hdisk27  87.1    1.6K   175.7  1.5K    64.1
hdisk7   71.1    29.7K  1.8K  17.4K   12.2K
hdisk5   64.1    52.2K  702.4  264.3   51.9K
hdisk8   58.6    29.2K  683.9  23.7K   5.5K
hdisk6   48.1    6.4K   661.3  262.3   6.2K
hdisk96  36.0    646.8  81.1   646.8   0.0
hdisk97  24.5    392.5  49.1   392.5   0.0
hdisk95  19.0    310.4  39.0   310.4   0.0
hdisk19  0.0     0.0    0.0    0.0     0.0
hdisk3   0.0     0.0    0.0    0.0     0.0

Name      PID  CPU%  PgSp  Owner
lrud      16392  17.9  0.8  root
sas.e9bb  2138252  17.8  8.1  pschuh2
spdsbase  1519652  17.0  3.6  ingxsas
sas.e9bb  1957970  12.9  6.9  egriff6
sas.e9bb  1601602  9.7   4.2  pgaliva
sas.e9bb  1998990  4.5   4.4  pgaliva
sas.e9bb  1642606  3.7   4.2  pgaliva
sas.e9bb  884866  3.3   7.4  egriff6
sas.e9bb  1941560  2.2   4.2  pgaliva
j2pg      159842  0.8   3.5  root
spdsbase  1044718  0.7  161.0  ingxsas
sas.e9bb  2244740  0.5   6.1  kbina

EVENTS/QUEUES  FILE/TTY
Cswitch  10236  Readch  202.5M
Syscall  72609  Writech  230.4M
Reads    17562  Rawin   0
Writes   14215  Ttyout  565
Forks    0      Igets   0
Execs    0      Namei   141
Runqueue 12.5   Dirblk  0
Waitqueue 2.5

PAGING
Faults    1690  MEMORY
Steals    51546  Real,MB  16384
PgspIn    0      $ Comp   42.2
PgspOut   0      $ Noncomp  57.7
PageIn    12075  $ Client  57.7
PageOut   46166
Sios      57980  PAGING SPACE
Size,MB   16384
$ Used    0.0
$ Free    100.0

NFS (calls/sec)
ServerV2  0
ClientV2  0  Press:
ServerV3  0  "h" for help
ClientV3  0  "q" to quit
    
```



Application

□ Fulltimer option

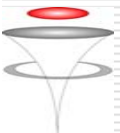
```
NOTE: PROCEDURE SQL used (Total process time):
real time          2:17:07.78
user cpu time      2:16.25
system cpu time    5:04.29
Memory             1263k
Page Faults       52924
Page Reclaims     27083
Page Swaps        0
Voluntary Context Switches 10078741
Involuntary Context Switches 359586
Block Input Operations 0
Block Output Operations 0
```

Overall SAS System	
real time (Secs)	6248
user cpu time (Secs)	281
system cpu time (Secs)	19
Memory(K)	16511
Page Faults	274
Page Reclaims	20453
Page Swaps	0
Voluntary Context Switches	239726
Involuntary Context Switches	30296
Block Input Operations	0
Block Output Operations	0

	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
	Data Step to Create all Records		Proc Format		Proc Sort by Character Variables		Proc Summary by HS codes all numeric variables		Proc Summary by HS10 QTY1 QTY2 STATMOYR		PROC SQL to create indexes		Real CPU	
Thread 1	148.91	125	0.3	0.2	81.03	19	578.97	49	594.84	48	798.47	214	4273	1013
Thread 2	149.14	129	0.45	0.37	138.48	20	625.54	53	646.71	58	857.73	327	4503	1450
Thread 3	161.23	142	0.35	0.35	251.97	113	590.6	341	669.97	362	858.82	154	4533	2150
Thread 4	154.54	143	0.38	0.42	232.72	118	559.05	378	674.72	353	847.02	88	4507	2036
Thread 5	161.34	142	0.38	0.34	290.47	115	550.58	395	680.73	360	849.09	103	4535	2065
Thread 6	174.45	145	0.44	0.31	292.73	116	545.84	394	688.72	372	844.47	60	4531	2045
Thread 7	174.35	144	0.51	0.4	301.42	128	537.78	363	732.94	370	864.62	121	4561	2069
Thread 8	170.57	145	0.47	0.28	295.81	146	535.62	355	694.33	372	871.05	51	4522	1998
Thread 9	164.05	147	0.56	0.32	324.12	133.32	655.75	365	988.41	327	742.44	66	5366	1867
Thread 10	167.28	149	0.96	0.37	307.01	123	694.14	344	1087.96	322	699.74	46	5430	1807
Thread 11	179	153	0.83	0.42	236.49	125	798.43	315	897.97	344	686.95	39	5070	1795
Thread 12	188.11	156	1.02	0.39	309	140	1013	416	1075	335	585	25	4961	1872
Thread 13	185	159	1.21	0.7	482	111	1048	394	935	340	479	22	4872	1812
Thread 14	190	160	1.09	0.63	442	144	1003	409	1012	342	500	22	4810	1831
Thread 15	204	162	1.41	0.58	573	146	1142	399	874	335	349	26	4705	1792
Thread 16	191	162	1.46	0.3	665	178	1270	400	831	332	304	17	4627	1761
Thread 17	194	166	1.9	0.91	426	189	742	423	628	332	512	20	4156	1754
Thread 18	188	174	1.81	0.71	473	180	695	400	616	330	520	21	4076	1709

Recommendations - Hardware

- ❑ Hard disk striping – Use both hardware controller and software striping
- ❑ RAID0 or RAID10 for WORK and UTILLOC file systems
 - http://www.acnc.com/04_01_10.html
- ❑ 4Gigabit fiber channels recommended
 - <http://www.fibrechannel.org/OVERVIEW/hardware.html>



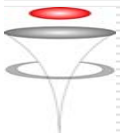
Recommendations - OS

□ WORK

- Use a separate file system other than /var
- Size of work file system should be 2 to 3 times bigger than your largest job size

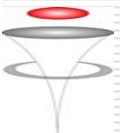
□ UTILLOC

- Create smaller file systems to use with UTILLOC. The size of the all the file systems should be equal or greater than the combined size of the datasets which use UTILLOC
- Number of file systems to be used should equal to the number of CPU count
- Checking cpu count
 - `bash$ lparstat`
 - System configuration: `type=Shared mode=Uncapped smt=On lcpu=8 mem=16384 psize=8 ent=2.00`
 - `%user %sys %wait %idle physc %entc lbusy app vcsw phint`
 - -----
 - 48.5 48.8 0.3 2.4 0.27 13.7 10.6 6.39 6403252110 505482495
- Summary, Sort, DMREG use UTILLOC



Recommendations - OS

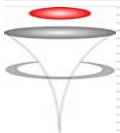
- ❑ Use different fiber channels for WORK and UTILLOC file system usage other than the data
- ❑ 2TB max size for a file system
- ❑ Paging space (Virtual Memory) should be at least equal to physical RAM and can be 2 to 3 times when possible



Recommendations – OS

- IBM recommended parameters for AIX

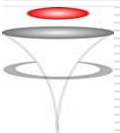
VMO Parameters	SASServer1	SASServer2	IBM recommendation
lru_file_repage	0	1	0
strict_maxclient	1	0	0
strict_maxperm	0	0	0
nokilluid	0	0	1
minperm%	10	15	5
maxclient%	80	30	80
maxperm%	80	30	80
minfree	960	960	1440 for SASServer1 and 960 for SASServer2



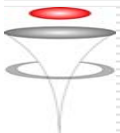
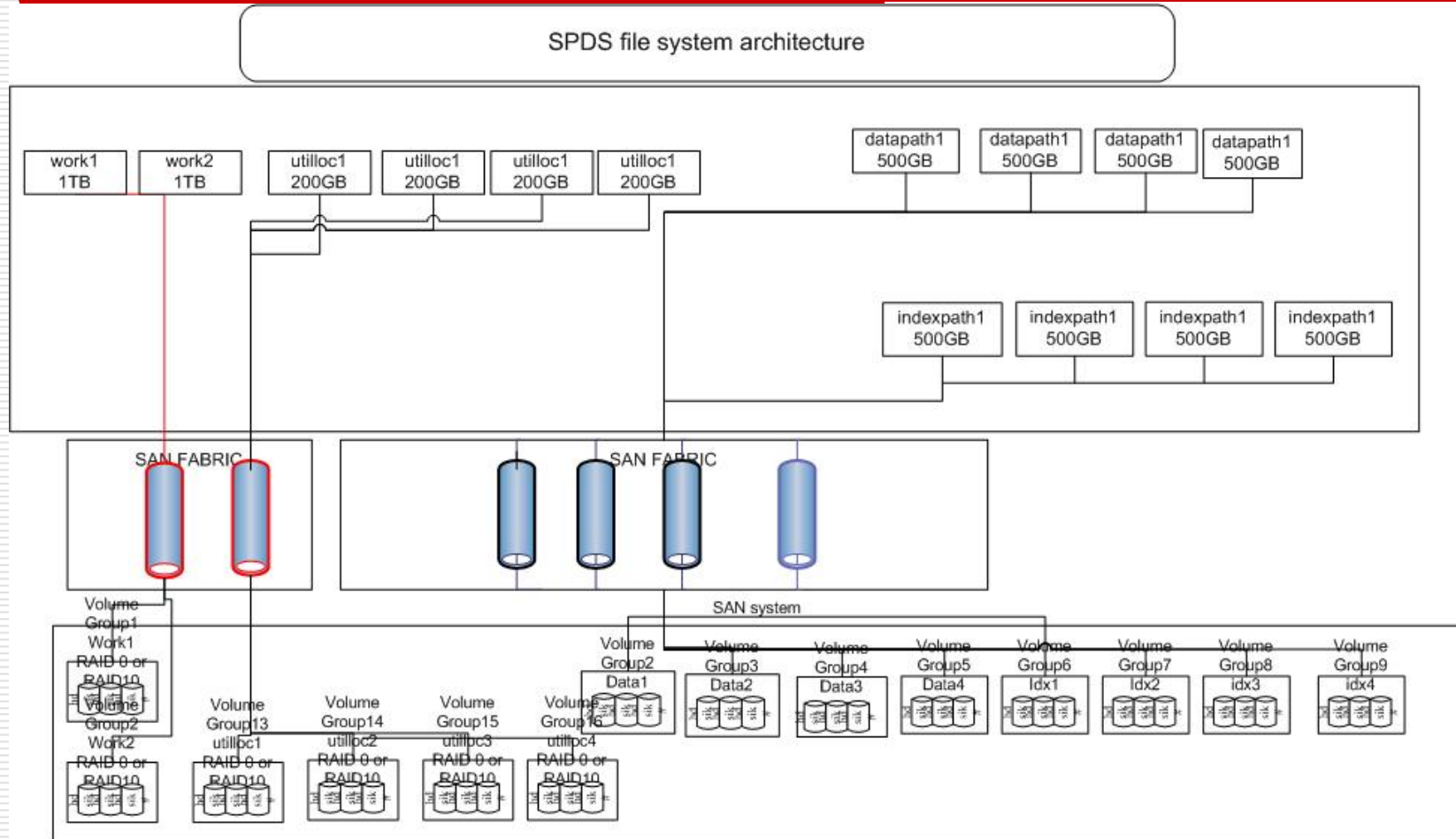
Recommendations - Application

□ For SPDS

- Use data paths, index paths and metadata paths in SPDSserver parm and libname parm files
- The number of data paths should be equal to the number of CPU count
- Set maxwhthreads to CPU count
- Use tmpdomain parm in spdserver parm file
- The minpartsize should be calculated to fit the most job sizes.
- To eliminate the traffic over the network for certain SAS procs, use client on the same system where SPDS Server is running.

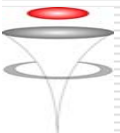


Recommendations – SPDS File System Structure



References and further reading

- ❑ http://www.sas.com/partners/directory/sun/sas9_on_solaris10_superheroes_paper.pdf
- ❑ <http://www.sas.com/partners/directory/sun/v9on9.pdf>
- ❑ <http://www.sas.com/partners/directory/sun/sugi28.pdf>
- ❑ <http://www.sas.com/partners/directory/sun/performance/index.html>
- ❑ <http://www.sas.com/partners/directory/ibm/AIXTuningGuide.pdf>



Questions?

□ Contact Basha.Shaik@gmail.com

