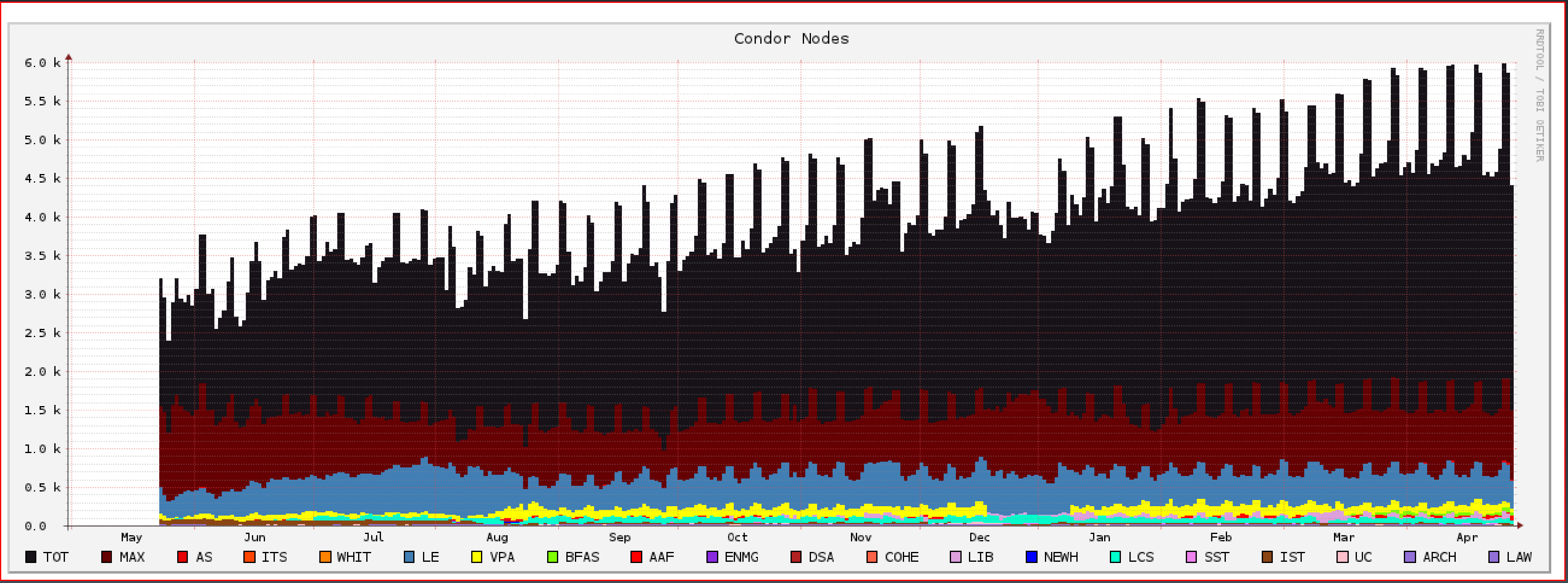
Managing a growing campus pool

Eric Sedore

essedore@syr.edu

Quick context

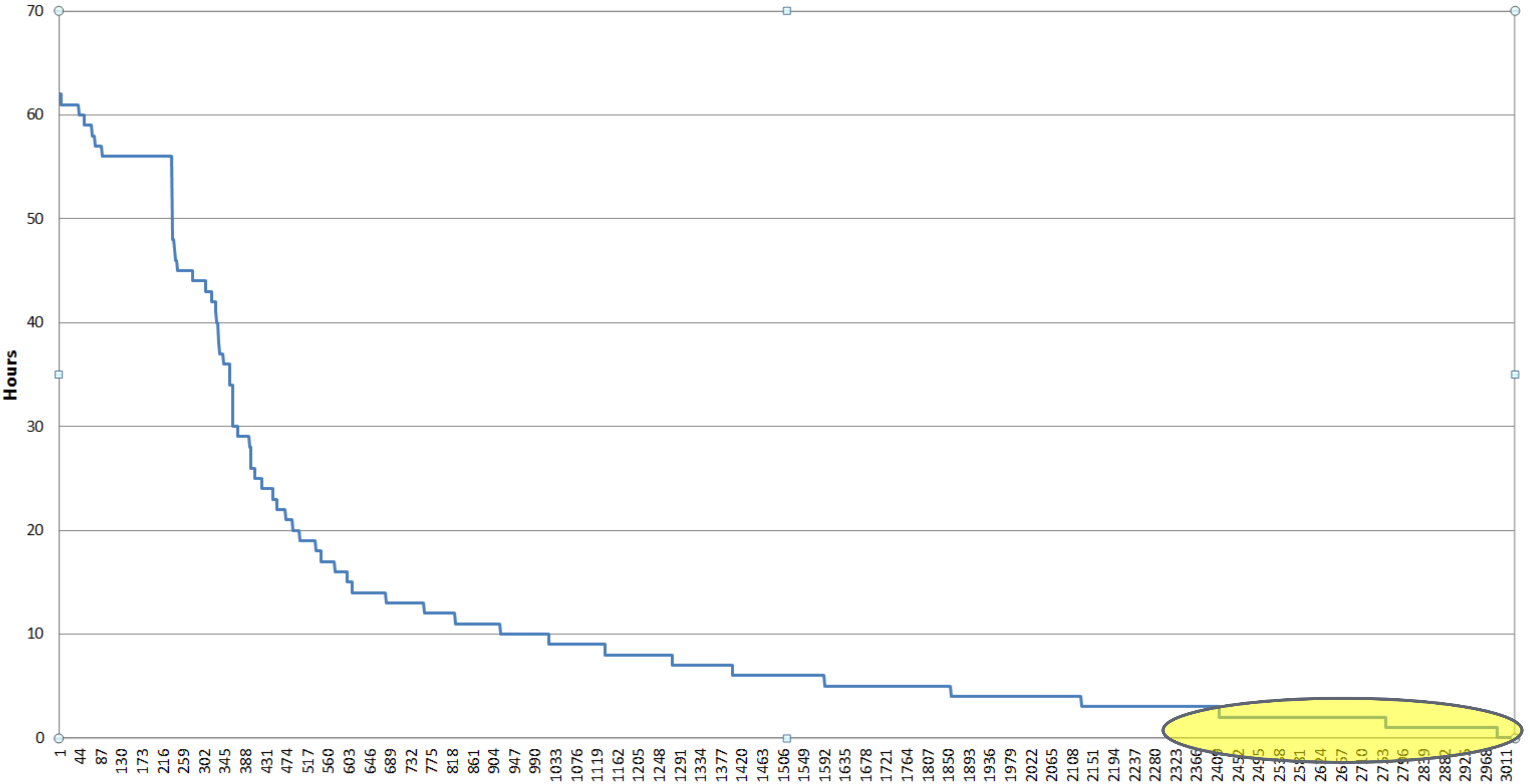
- Scavenging CPU time from 2000 desktops on campus
- Growth from 3000 to 6000+ cores over the last year
- Varied types of research (in all dimensions, run time, data needed per job, data access type (HTCondor transfer/NFS))
- Challenge feeding data to the compute nodes
  - VPN tunnel bandwidth
  - SAN Storage / Virtual server environment
  - Campus network
- Exploring steering jobs via average availability of each node


Condor Nodes
6.0 k
5.5 k
5.0 k
4.5 k
4.0 k
3.5 k
3.0 k
2.5 k
2.0 k
1.5 k
1.0 k
0.5 k
0.0
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
TOT
MAX
AS
ITS
WHIT
LE
VPA
BFAS
AAF
ENMG
DSA
COHE
LIB
NEWH
LCS
SST
IST
UC
ARCH
LAW
RRDTOOL / TOBI OETIKER

SYRACUSE UNIVERSITY ITS Information Technology and Services

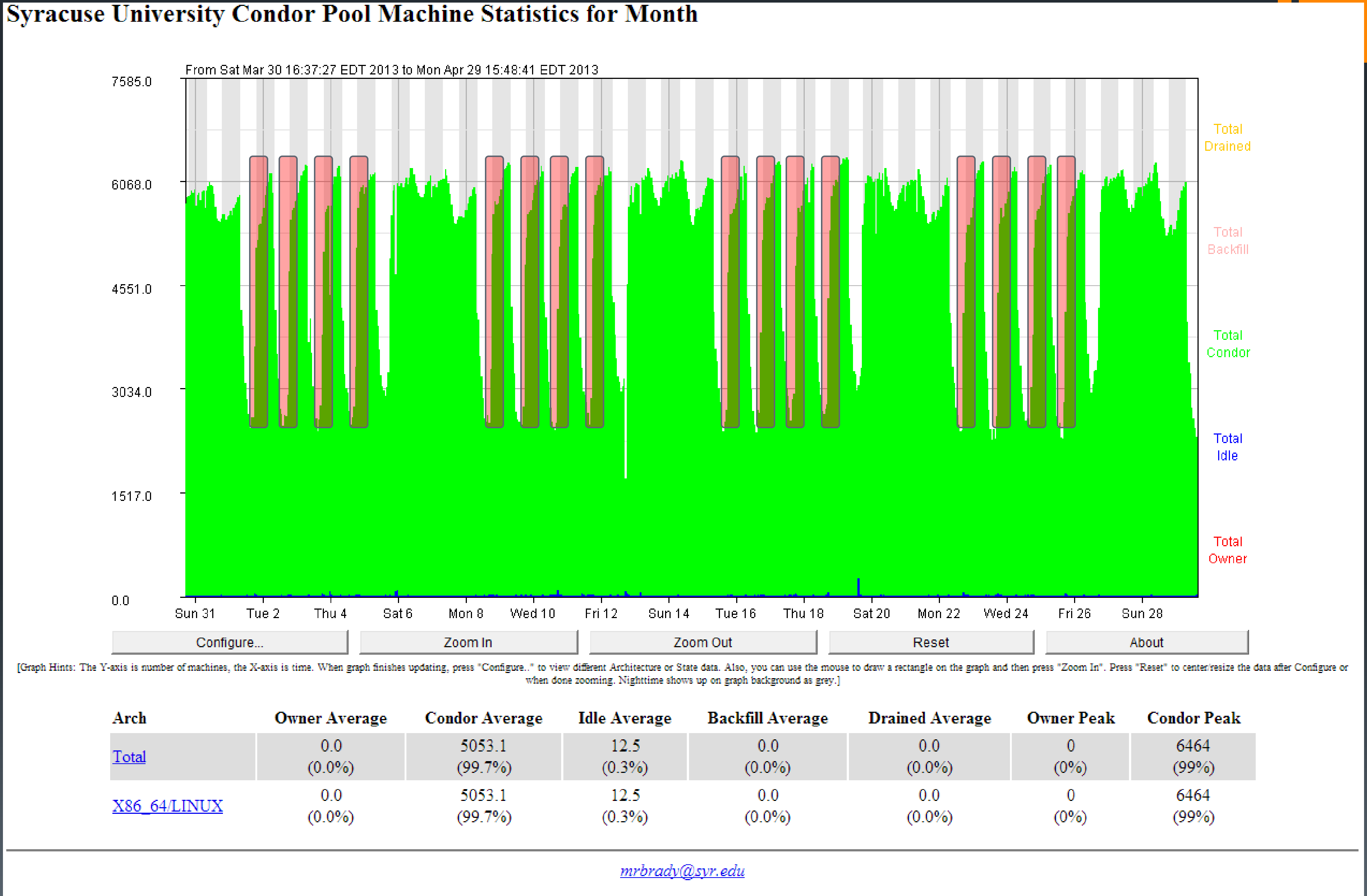
Steering jobs

- More Traditional – CPU/Memory/Application requirements
- Jobs that are not easily checkpointed and have longer run times
- Dynamic average giving the job a view into the amount of scavenging time (in hours) of a node – publish via classad
  - sessionlength = "44"
  - sessionlength = "9"
  - sessionlength = "10"
  - sessionlength = "41"
- Add to job requirements or ranking: TARGET.sessionlength >= "4“
- Data gathered through parsing logs from the last seven days (from Condor VM Coordinator)


Hours
70
60
50
40
30
20
10
0
1
44
87
130
173
216
259
302
345
388
431
474
517
560
603
646
689
732
775
818
861
904
947
990
1033
1076
1119
1162
1205
1248
1291
1334
1377
1420
1463
1506
1549
1592
1635
1678
1721
1764
1807
1850
1893
1936
1979
2022
2065
2108
2151
2194
2237
2280
2323
2366
2409
2452
2495
2538
2581
2624
2667
2710
2753
2796
2839
2882
2925
2968
3011

Dealing with growth

- Data delivery infrastructure - horizontal and vertical scaling
- Network infrastructure upgrades – upgrading 1 Gb links to 10 Gb links
- Scaling out VPN end points – working toward a model, 1000-1500 per endpoint
- Throttling vs. bottleneck (protecting network links where there are high compute node populations)



Syracuse University Condor Pool Machine Statistics for Month

[Graph Hints: The Y-axis is number of machines, the X-axis is time. When graph finishes updating, press "Configure.." to view different Architecture or State data. Also, you can use the mouse to draw a rectangle on the graph and then press "Zoom In". Press "Reset" to center/resize the data after Configure or when done zooming. Nighttime shows up on graph background as grey.]

Arch	Owner Average	Condor Average	Idle Average	Backfill Average	Drained Average	Owner Peak	Condor Peak
Total	0.0 (0.0%)	5053.1 (99.7%)	12.5 (0.3%)	0.0 (0.0%)	0.0 (0.0%)	0 (0%)	6464 (99%)
X86_64/LINUX	0.0 (0.0%)	5053.1 (99.7%)	12.5 (0.3%)	0.0 (0.0%)	0.0 (0.0%)	0 (0%)	6464 (99%)

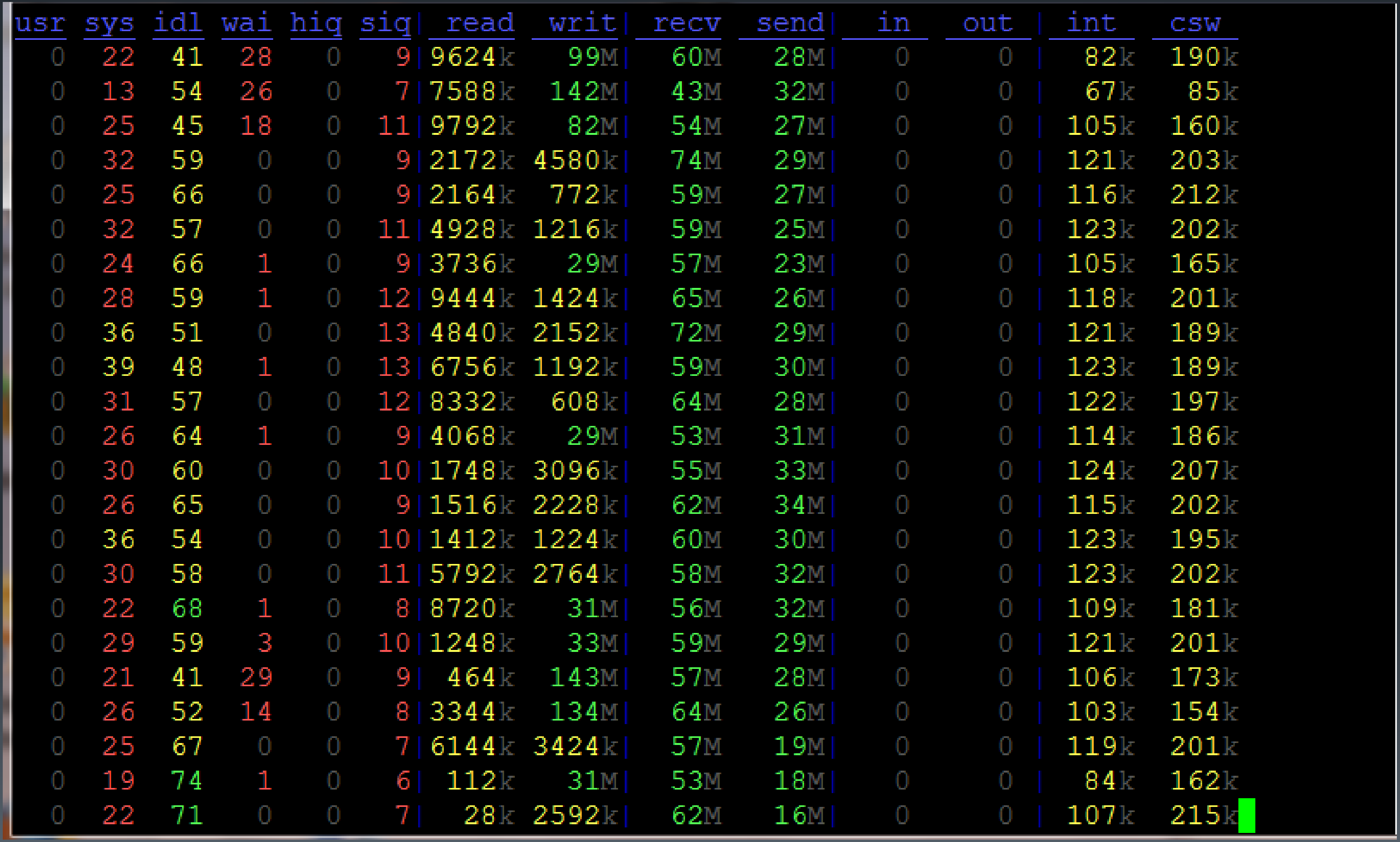
mrbrady@syr.edu

Pushing NFS

- ~6300 nodes supported on a single NFS server (65000+ open files)
- Can run out of CPU at peak times (current server has 16 cores) – context switching / I/O wait
- Virtual Machine running in ESX 5.x – SAN storage (15K spindles on an IBM 5300)

- 3.0.0-21-virtual #35-Ubuntu SMP Fri May 25 18:35:12 UTC 2012 x86_64 x86_64 x86_64 GNU/Linux

Pool up and running – supporting the load



```
usr sys idl wai hiq siq| read  writ| recv  send|  in   out | int   csw
  0  22  41  28   0   9|9624k   99M|  60M   28M|   0     0 |  82k  190k
  0  13  54  26   0   7|7588k  142M|  43M   32M|   0     0 |  67k   85k
  0  25  45  18   0  11|9792k   82M|  54M   27M|   0     0 | 105k  160k
  0  32  59   0   0   9|2172k 4580k|  74M   29M|   0     0 | 121k  203k
  0  25  66   0   0   9|2164k  772k|  59M   27M|   0     0 | 116k  212k
  0  32  57   0   0  11|4928k 1216k|  59M   25M|   0     0 | 123k  202k
  0  24  66   1   0   9|3736k   29M|  57M   23M|   0     0 | 105k  165k
  0  28  59   1   0  12|9444k 1424k|  65M   26M|   0     0 | 118k  201k
  0  36  51   0   0  13|4840k 2152k|  72M   29M|   0     0 | 121k  189k
  0  39  48   1   0  13|6756k 1192k|  59M   30M|   0     0 | 123k  189k
  0  31  57   0   0  12|8332k  608k|  64M   28M|   0     0 | 122k  197k
  0  26  64   1   0   9|4068k   29M|  53M   31M|   0     0 | 114k  186k
  0  30  60   0   0  10|1748k 3096k|  55M   33M|   0     0 | 124k  207k
  0  26  65   0   0   9|1516k 2228k|  62M   34M|   0     0 | 115k  202k
  0  36  54   0   0  10|1412k 1224k|  60M   30M|   0     0 | 123k  195k
  0  30  58   0   0  11|5792k 2764k|  58M   32M|   0     0 | 123k  202k
  0  22  68   1   0   8|8720k   31M|  56M   32M|   0     0 | 109k  181k
  0  29  59   3   0  10|1248k   33M|  59M   29M|   0     0 | 121k  201k
  0  21  41  29   0   9| 464k  143M|  57M   28M|   0     0 | 106k  173k
  0  26  52  14   0   8|3344k  134M|  64M   26M|   0     0 | 103k  154k
  0  25  67   0   0   7|6144k 3424k|  57M   19M|   0     0 | 119k  201k
  0  19  74   1   0   6| 112k   31M|  53M   18M|   0     0 |  84k  162k
  0  22  71   0   0   7|  28k 2592k|  62M   16M|   0     0 | 107k  215k
```