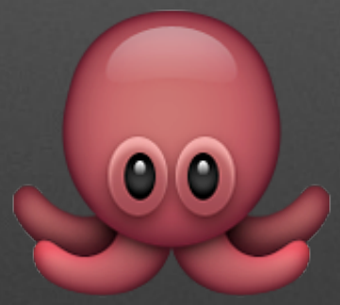
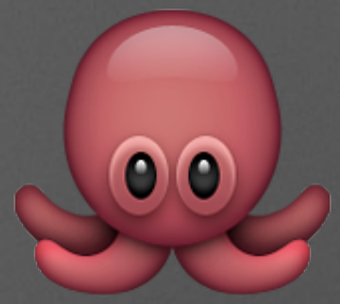
Monitoring jobs

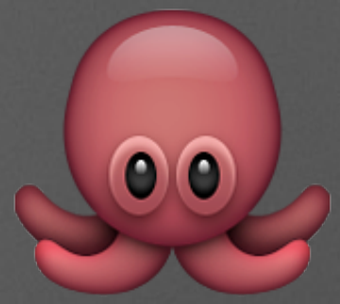


Monitoring jobs



CPU	RAM	Disk	Jobs
top	top	df	qstat
ps	vmstat	du	qconf
sar	free		qacct
			qdel
			jobs

Monitoring jobs: CPU



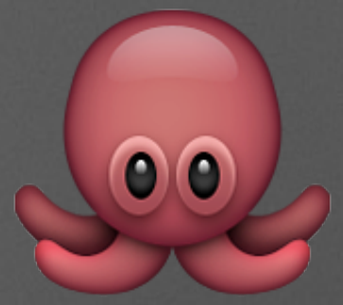
`top`: continuous look at processes

rdikow — ssh — 89×35

```
top - 15:51:02 up 6 days, 22:58,  4 users,  load average: 0.03, 0.13, 0.16
Tasks: 615 total,   1 running, 614 sleeping,   0 stopped,   0 zombie
Cpu(s):  0.0%us,  0.1%sy,  0.0%ni, 99.9%id,  0.0%wa,  0.0%hi,  0.0%si,  0.0%st
Mem:  65957020k total,  5912796k used, 60044224k free,   890452k buffers
Swap: 51199996k total,        0k used, 51199996k free,  3309204k cached

  PID USER      PR  NI  VIRT  RES  SHR S %CPU %MEM    TIME+  COMMAND
 3423 root      20   0  420m 231m 9.9m S  0.7  0.4  27:26.44 splunkd
51345 dikowr    20   0 15436 1636  896 R  0.7  0.0   0:00.50 top
  104 root      20   0     0    0    0 S  0.3  0.0   0:16.01 events/5
 2374 root      20   0 18372  812  468 S  0.3  0.0   5:00.50 irqbalance
34494 root      20   0     0    0    0 S  0.3  0.0   1:30.25 nfsd
    1 root      20   0 19348 1504 1180 S  0.0  0.0   0:01.80 init
    2 root      20   0     0    0    0 S  0.0  0.0   0:00.03 kthreadd
    3 root      RT   0     0    0    0 S  0.0  0.0   0:00.09 migration/0
    4 root      20   0     0    0    0 S  0.0  0.0   0:02.81 ksoftirqd/0
    5 root      RT   0     0    0    0 S  0.0  0.0   0:00.00 stopper/0
    6 root      RT   0     0    0    0 S  0.0  0.0   0:00.46 watchdog/0
    7 root      RT   0     0    0    0 S  0.0  0.0   0:00.14 migration/1
    8 root      RT   0     0    0    0 S  0.0  0.0   0:00.00 stopper/1
    9 root      20   0     0    0    0 S  0.0  0.0   0:00.37 ksoftirqd/1
   10 root      RT   0     0    0    0 S  0.0  0.0   0:00.46 watchdog/1
   11 root      RT   0     0    0    0 S  0.0  0.0   0:05.95 migration/2
   12 root      RT   0     0    0    0 S  0.0  0.0   0:00.00 stopper/2
   13 root      20   0     0    0    0 S  0.0  0.0   0:01.93 ksoftirqd/2
   14 root      RT   0     0    0    0 S  0.0  0.0   0:00.44 watchdog/2
   15 root      RT   0     0    0    0 S  0.0  0.0   0:06.76 migration/3
   16 root      RT   0     0    0    0 S  0.0  0.0   0:00.00 stopper/3
   17 root      20   0     0    0    0 S  0.0  0.0   0:02.08 ksoftirqd/3
   18 root      RT   0     0    0    0 S  0.0  0.0   0:00.44 watchdog/3
   19 root      RT   0     0    0    0 S  0.0  0.0   0:00.33 migration/4
   20 root      RT   0     0    0    0 S  0.0  0.0   0:00.00 stopper/4
   21 root      20   0     0    0    0 S  0.0  0.0   0:03.63 ksoftirqd/4
   22 root      RT   0     0    0    0 S  0.0  0.0   0:00.40 watchdog/4
   23 root      RT   0     0    0    0 S  0.0  0.0   0:00.31 migration/5
```

`ps`: snapshot of processes

Monitoring jobs: CPU



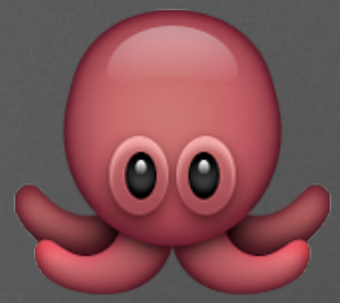
`sar`: System Activity Report, part of Sysstat

SAR.txt

```
06:00:01 PM     all     22.09      0.00      0.09      0.00      0.00     77.82
06:10:01 PM     all     22.23      0.00      0.01      0.00      0.00     77.76
06:20:01 PM     all     22.09      0.00      0.09      0.00      0.00     77.82

06:20:01 PM     CPU     %user     %nice   %system   %iowait    %steal     %idle
06:30:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.76
06:40:01 PM     all     22.08      0.00      0.09      0.00      0.00     77.83
06:50:01 PM     all     22.18      0.00      0.02      0.00      0.00     77.80
07:00:01 PM     all     22.14      0.00      0.08      0.00      0.00     77.78
07:10:01 PM     all     22.08      0.00      0.09      0.00      0.00     77.83
07:20:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.76
07:30:01 PM     all     13.38      0.00      0.01      0.00      0.00     86.61
07:40:01 PM     all      0.69      0.00      0.01      0.00      0.00     99.30
07:50:01 PM     all      7.49      0.00      0.08      0.00      0.00     92.43
08:00:02 PM     all     22.21      0.00      0.03      0.00      0.00     77.75
08:10:01 PM     all     22.11      0.00      0.09      0.00      0.00     77.80
08:20:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.76
08:30:01 PM     all     22.09      0.00      0.09      0.00      0.00     77.82
08:40:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.75
08:50:01 PM     all     22.08      0.00      0.09      0.00      0.00     77.83
09:00:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.75
09:10:01 PM     all     22.09      0.00      0.09      0.00      0.00     77.82
09:20:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.76
09:30:01 PM     all     22.08      0.00      0.09      0.00      0.00     77.83
09:40:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.76
09:50:01 PM     all     22.09      0.00      0.09      0.00      0.00     77.81
10:00:01 PM     all     22.24      0.00      0.01      0.00      0.00     77.75

10:00:01 PM     CPU     %user     %nice   %system   %iowait    %steal     %idle
```

Can access for a particular day or set of days in the past month
`sar -b -f /var/log/sa/sa20` (for the 20th of the month)
syntax varies by system

Monitoring jobs: RAM



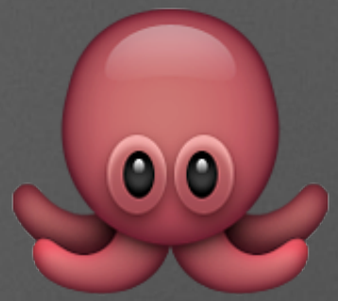
`vmstat:` get info on memory usage

```
dikowr@hydra-3% vmstat
procs -----------memory---------- ---swap-- -----io---- --system-- -----cpu-----
 r  b   swpd   free   buff  cache   si   so    bi    bo   in   cs us sy id wa st
 0  1      0 60039684 890464 3309216    0    0     1     1    3    3  0  0 99  1  0
```

`free:` similar to vmstat, output looks a bit different

```
[dikowr@login-3-1 ~]$ free
             total       used       free     shared    buffers     cached
Mem:      32998248    2193796   30804452        224     271480    1162592
-/+ buffers/cache:     759724   32238524
Swap:     31211516          0   31211516
```

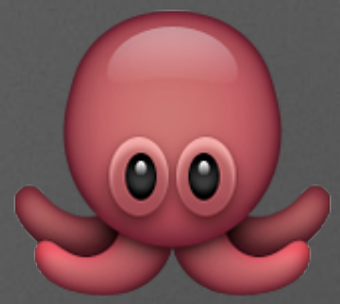
Monitoring jobs: disk



`df` disk free: `[-h]` display in KB, MB, or GB

`du` disk usage: `[-sh]` sum of directories

Monitoring jobs: Jobs



`qstat`

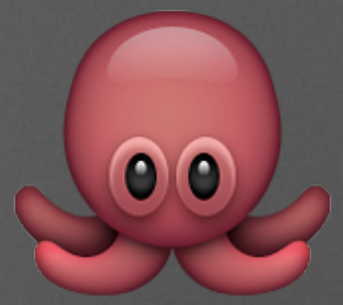
`qstat -f job_id` Produce detailed report about job

`qstat -u dikowR` To list all the jobs belonging to a particular user (me)

`qstat -q` To view a summary of all queues

`qstat -j` Prints info about pending jobs

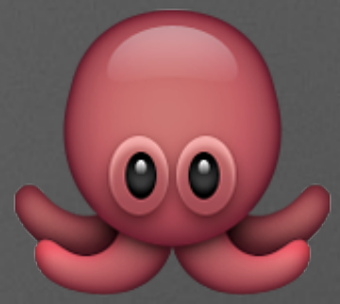
Monitoring jobs



Hydra website:

https://www.cfa.harvard.edu/~sylvain/hpc/status/

Monitoring jobs



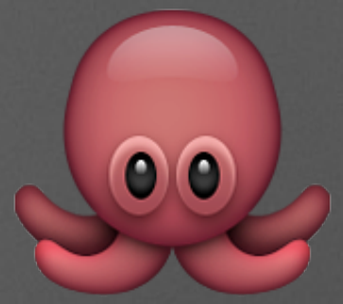
`qconf` get info on queue configuration

`qdel` delete a job (*e.g.* `$qdel job-id`)

`qacct` report on SGE usage

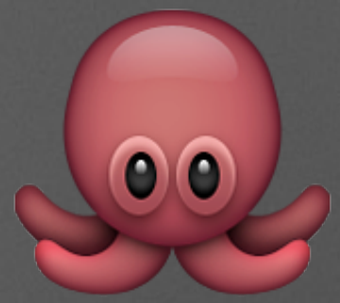
`jobs` display status of jobs in foreground/ background

Monitoring jobs



How to use shell scripting with Hydra to be productive in your job submission

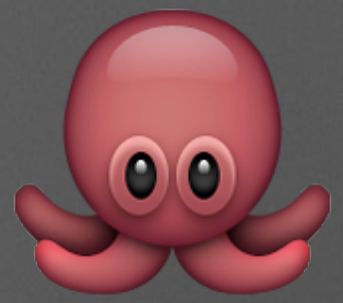
Monitoring jobs



```
#!/bin/bash

for i in *.nex; do

    qsub -q mThC.q -N MrBayes_$i -S /
bin/sh -cwd -j y –o MrBayes_$i.out –l
mres=1G,h_data=1G,h_vmem=1G mb.job $i

done
```

check out: `man qsub`

Monitoring jobs



```
mb $1
```

Compute intensive
RAM intensive
I/O intensive

Compute intensive

-Limitation is processing

-How to parallelize: decompose the overall application process into separate tasks, which can then be executed in parallel

Common examples:

Phylogenetic tree search

Alignment

I/O (data) intensive

-Spends more time reading and writing than processing

-Should scale linearly according to the size of the data

-How to parallelize: split the data into parts that can be processed independently and then re-assemble

Common examples:

BLAST

lots of little assemblies

RAM intensive

-Limitation is memory

-Will likely crash if not enough RAM, more time won't help

Common examples:

Genome/transcriptome assembly
(~300GB for 1Gbp genome)

Phylogenetic tree building on a very* large matrix
(# of taxa vs. # of characters)

Resources

Google

Biostars

Stack Overflow

Individual program email lists/google groups

Resources

https://hydra-3.si.edu/wordpress/

https://hydra-3.si.edu/ganglia/

http://si-vmatlassian:8090/display/ORIS/Migration+Notes%3A+Hydra-2+to+Hydra-3

http://si-vmatlassian:8090/display/ORIS/Hydra+Software

Please fill out survey of what you'd like to see next…