



Hidden Performance Tooling

Marc Skinner

What do mean by Performance?

- **Is my system slow?**
- **What is my system doing?**
- **Do I have a bottleneck?**
- How has my system been performing – trends?
- What could be the peak performance – benchmarking!
- Monitoring?
- Threshold alerting



Performance tooling in RHEL

- Tools we all use today
 - top
 - vmstat
 - sar
 - iostat
 - free
 - snmp
 - ethtool
 - tuned-adm



Hidden tooling in RHEL7

- Networking
- CPU
- Disk
- Multi-functional



iptraf-ng [NETWORKING]

- TUI to analyze
 - Network performance
 - Statics
 - Flows
- Quick tool to use

iptraf-ng 1.1.4

```
IP traffic monitor
General interface statistics
Detailed interface statistics
Statistical breakdowns...
LAN station monitor
```

```
Filters...
```

```
Configure...
```

```
About...
```

```
Exit
```

Displays more statistics for a selected interface

Up/Down-Move selector Enter-execute

nc [NETWORKING]

- nc = NetCat (RHEL6 = nc, RHEL7 = nmap-ncat)
- Useful for testing bandwidth performance of 2 devices – point A to point B

- **Node 1 (point A)**

```
# nc -l -n 12345 > /dev/null
```

- **Node 2 (point B :: 1gb NIC)**

```
# dd if=/dev/zero bs=1M count=1024 | nc -n 192.168.33.13 12345
```

```
1024+0 records in
```

```
1024+0 records out
```

```
1073741824 bytes (1.1 GB) copied, 53.3481 s, 20.1 MB/s
```

- **Node 3 (point C :: 1gb NIC)**

```
#dd if=/dev/zero bs=1M count=1024 | nc -n 192.168.33.44 12345
```

```
1024+0 records in
```

```
1024+0 records out
```

```
1073741824 bytes (1.1 GB) copied, 9.11378 s, 118 MB/s
```



qperf [NETWORKING]

- Tool to measure bandwidth and latency between two nodes
- **Node 1 (point A :: 1gb NIC)**

qperf

- **Node 2 (point B :: 1gb NIC)**

qperf 192.168.33.13 tcp_bw tcp_lat

tcp_bw:

bw = **19.2** MB/sec

tcp_lat:

latency = 276 us

qperf 192.168.33.13 udp_bw udp_lat

udp_bw:

send_bw = **120** MB/sec

recv_bw = 78.2 MB/sec

udp_lat:

latency = 552 us



ibstatus [NETWORKING]

- Infiniband interface status – similar to ethtool for Ethernet interfaces

```
# ibstatus
```

```
Infiniband device 'mthca0' port 1 status:
```

```
default gid:   fe80:0000:0000:0000:0005:ad00:000c:6de1
```

```
base lid: 0x4
```

```
sm lid:       0x1
```

```
state:        4: ACTIVE
```

```
phys state:   5: LinkUp
```

```
rate:         20 Gb/sec (4X DDR)
```

```
link_layer:   InfiniBand
```



ifstat [NETWORKING]

- Dumps networking statistics

```
[root@iscsi ~]# ifstat
#kernel
Interface      RX Pkts/Rate    TX Pkts/Rate    RX Data/Rate    TX Data/Rate
              RX Errs/Drop   TX Errs/Drop   RX Over/Rate    TX Coll/Rate
lo              0 0             0 0             0 0             0 0
              0 0             0 0             0 0             0 0
enp8s0          0 0             0 0             0 0             0 0
              0 0             0 0             0 0             0 0
enp2s0f0        158 0           57 0            19264 0         10879 0
              0 0             0 0             0 0             0 0
enp2s0f1        0 0             0 0             0 0             0 0
              0 0             0 0             0 0             0 0
enp9s0          326 0           6 0             32012 0         640 0
              0 0             0 0             0 0             0 0
enp10s0         198 0           6 0             25312 0         670 0
              0 0             0 0             0 0             0 0
ib0             10572 0         13372 0         4909K 0         131868K 0
              0 14            0 0             0 0             0 0
ib1             14 0            0 0             1232 0          0 0
              0 14            0 0             0 0             0 0
team0           682 0           55 0            64312 0         10239 0
              0 0             0 0             0 0             0 0
[root@iscsi ~]#
```



htop [CPU]

- TUI for running processes – lets you scroll horizontally and vertically
- Similar to top

```

1 [ | 0.7%] 4 [ 0.0%]
2 [ | 0.0%] 5 [ 0.0%]
3 [ | 0.7%] 6 [ 0.0%]
Mem[||||] 953M/31.3G Tasks: 45, 24 thr; 1 running
Swp[ ] 0K/8.00G Load average: 0.00 0.01 0.05
Uptime: 47 days, 01:19:34

PID USER PRI NI VIRT RES SHR S CPU% MEM% TIME+ Command
3048 root 20 0 240M 4364 1128 S 0.0 0.0 18:30.78 /usr/bin/hptsrv
30655 root 20 0 120M 3344 1476 R 0.0 0.0 0:00.32 htop
30192 root 20 0 142M 5672 3896 S 0.0 0.0 0:00.37 sshd: root@pts/1
2573 root 20 0 251M 19016 6736 S 0.0 0.1 5:16.11 targetd
2687 root 20 0 240M 4364 1128 S 0.0 0.0 18:37.02 /usr/bin/hptsrv
1987 root 20 0 33404 1928 1516 S 0.0 0.0 11:06.12 /usr/bin/teamd -U -D -o -t teamd
29096 root 20 0 141M 5144 3876 S 0.0 0.0 0:00.17 sshd: root@pts/0
1 root 20 0 193M 13272 3936 S 0.0 0.0 2:33.94 /usr/lib/systemd/systemd --swi
704 root 20 0 104M 52808 48296 S 0.0 0.2 0:35.67 /usr/lib/systemd/systemd-journ
2713 root 20 0 406M 8324 2628 S 0.0 0.0 0:00.00 /usr/sbin/lvmetad -f
718 root 20 0 406M 8324 2628 S 0.0 0.0 0:00.25 /usr/sbin/lvmetad -f
730 root 20 0 45888 4712 2776 S 0.0 0.0 0:01.21 /usr/lib/systemd/systemd-udev
1182 root 16 -4 113M 1632 1236 S 0.0 0.0 0:02.71 /sbin/auditd -n
1179 root 16 -4 113M 1632 1236 S 0.0 0.0 0:06.29 /sbin/auditd -n
1207 root 20 0 26396 1764 1436 S 0.0 0.0 0:39.09 /usr/lib/systemd/systemd-logind
1208 root 20 0 19284 1208 960 S 0.0 0.0 2:40.41 /usr/sbin/irqbalance --foregro
1220 root 20 0 318M 26464 25692 S 0.0 0.1 0:05.21 /usr/sbin/rsyslogd -n
1221 root 20 0 318M 26464 25692 S 0.0 0.1 0:03.05 /usr/sbin/rsyslogd -n
1211 root 20 0 318M 26464 25692 S 0.0 0.1 0:08.28 /usr/sbin/rsyslogd -n
1222 dbus 20 0 37160 2188 1460 S 0.0 0.0 0:00.00 /bin/dbus-daemon --system --ad
1212 dbus 20 0 37160 2188 1460 S 0.0 0.0 1:10.16 /bin/dbus-daemon --system --ad
1214 root 20 0 198M 1252 796 S 0.0 0.0 0:00.00 /usr/sbin/gssproxy -D
1215 root 20 0 198M 1252 796 S 0.0 0.0 0:00.00 /usr/sbin/gssproxy -D
1216 root 20 0 198M 1252 796 S 0.0 0.0 0:00.00 /usr/sbin/gssproxy -D
1217 root 20 0 198M 1252 796 S 0.0 0.0 0:00.00 /usr/sbin/gssproxy -D
1218 root 20 0 198M 1252 796 S 0.0 0.0 0:00.00 /usr/sbin/gssproxy -D
1213 root 20 0 198M 1252 796 S 0.0 0.0 0:02.82 /usr/sbin/gssproxy -D
1223 root 20 0 26928 2656 1728 S 0.0 0.0 0:06.33 /usr/sbin/smartd -n -q never
1226 libstorag 20 0 8528 820 672 S 0.0 0.0 0:05.18 /usr/bin/lsmc -d
1228 root 20 0 205M 5372 3656 S 0.0 0.0 0:00.00 /usr/sbin/abrt -d -s
1231 root 20 0 203M 4508 3156 S 0.0 0.0 0:02.46 /usr/bin/abrt-watch-log -F BUG
1232 root 20 0 5308 1004 536 S 0.0 0.0 0:01.75 /sbin/mdadm --monitor --scan -f
1259 root 20 0 426M 7848 6104 S 0.0 0.0 0:00.00 /usr/sbin/NetworkManager --no-d
1284 root 20 0 426M 7848 6104 S 0.0 0.0 0:10.41 /usr/sbin/NetworkManager --no-d
1233 root 20 0 426M 7848 6104 S 0.0 0.0 0:37.14 /usr/sbin/NetworkManager --no-d
1249 chrony 20 0 113M 1896 1504 S 0.0 0.0 0:04.43 /usr/sbin/chronyd
1293 root 20 0 53060 2688 2100 S 0.0 0.0 0:00.00 /usr/sbin/wpa supplicant -u -f
1295 polkitd 20 0 517M 14176 4676 S 0.0 0.0 0:00.00 /usr/lib/polkit-1/polkitd --no
1296 polkitd 20 0 517M 14176 4676 S 0.0 0.0 0:09.93 /usr/lib/polkit-1/polkitd --no
1297 polkitd 20 0 517M 14176 4676 S 0.0 0.0 0:00.00 /usr/lib/polkit-1/polkitd --no
1298 polkitd 20 0 517M 14176 4676 S 0.0 0.0 0:00.00 /usr/lib/polkit-1/polkitd --no
F1Help F2Setup F3Search F4Filter F5Tree F6SortBy F7Nice F8Nice +F9Kill F10Quit

```

virt-top [CPU]

- Top like tool for virtualized resources when using KVM

```
virt-top 11:28:26 - x86 64 8/8CPU 3325MHz 32000MB
6 domains, 6 active, 6 running, 0 sleeping, 0 paused, 0 inactive D:0 0:0 X:0
CPU: 3.6% Mem: 45056 MB (45056 MB by guests)
```

ID	S	RDRQ	WRRQ	RXBY	TXBY	%CPU	%MEM	TIME	NAME
2	R	0	0	4735	55K	1.7	25.0	185:33:11	WIN7
3	R	0	14	264	264	0.8	25.0	8:50.51	OSE3-MASTER1
5	R	0	0	132	132	0.6	25.0	8:04.34	OSE3-NODE1
6	R	0	0	188	174	0.3	25.0	4:03.07	OSE3-NODE2
7	R	0	32	9136	0	0.1	12.0	1:23.63	CALAMARI
4	R	0	0	0	0	0.0	25.0	1:26.95	OSE3-MASTER2



dstat [DISK]

- Great tool for viewing vmstat, iostat, ifstat interactively.
- Can build your table by including/excluding many options
- Nice for correlation of multiple resources

```
[root@iscsi ~]# dstat
You did not select any stats, using -cdngy by default.
---total-cpu-usage--- -dsk/total- -net/total- ---paging-- ---system--
usr  sys  idl  wai  hiq  sig| read  writ| recv  send|  in  out|  int  csw
13   0   86   0   0   0| 25M   908k|   0    0|   0   0| 1782  906
0   0  100   0   0   0| 2481k 29k|  24k 2500k|   0   0|  657  564
0   0  100   0   0   0| 2505k 74k|  48k 2522k|   0   0|  720  636
0   0  100   0   0   0| 1000k 46k|  18k 1005k|   0   0|  294  323
0   0  100   0   0   0|   0   127k|  64k 3888B|   0   0|  267  342
0   0  100   0   0   0| 2028k 2176k|  13k 1008k|   0   0|  298  270
0   0  100   0   0   0| 8192B 366k| 155k  11k|   0   0|  231  265
0   0  100   0   0   0| 1004k 31k|  13k 1008k|   0   0|  205  203
0   0  100   0   0   0|   0   43k|  23k 2208B|   0   0|  157  224
0   0  100   0   0   0| 1004k 96k|  47k 1010k|   0   0|  314  322
0   0  100   0   0   0|   0   46k|  22k 2208B|   0   0|  158  214
0   0  100   0   0   0| 1000k 29k|  13k 1004k|   0   0|  196  201
0   0  100   0   0   0|   0   90k|  23k 2208B|   0   0|  197  292
0   0  100   0   0   0| 2481k 29k|  24k 2498k|   0   0|  674  569
0   0  100   0   0   0| 2505k 116k|  69k 2525k|   0   0|  802  726
0   0  100   0   0   0| 2796k 3785k|  46k 1009k|   0   0|  407  406
0   0  100   0   0   0| 4096B 97k|  22k 6464B|   0   0|  191  243
```



iostat [DISK]

- Simple I/O top tool

```
Total DISK READ :    20.51 M/s | Total DISK WRITE :    0.00 B/s
Actual DISK READ:    20.51 M/s | Actual DISK WRITE:   19.56 M/s
```

TID	PRIO	USER	DISK READ	DISK WRITE	SWAPIN	IO>	COMMAND
29221	be/4	root	1001.27 K/s	0.00 B/s	0.00 %	0.00 %	[iscsi_trx]
25135	be/4	root	1752.22 K/s	0.00 B/s	0.00 %	0.00 %	[kworker/2:2]
28061	be/4	root	5.38 M/s	0.00 B/s	0.00 %	0.00 %	[kworker/3:1]
16299	be/4	root	9.29 M/s	0.00 B/s	0.00 %	0.00 %	[kworker/4:1]
27103	be/4	root	2.20 M/s	0.00 B/s	0.00 %	0.00 %	[kworker/0:2]
8676	be/4	root	977.80 K/s	0.00 B/s	0.00 %	0.00 %	[iscsi_trx]
1	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	systemd --s-erial
2	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[kthreadd]
3	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[ksoftirqd/0]
7	rt/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[migration/0]
8	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcu_bh]
9	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuob/0]
10	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuob/1]
11	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuob/2]
12	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuob/3]
13	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuob/4]
14	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuob/5]
15	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcu_sched]
16	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuos/0]
17	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuos/1]
18	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuos/2]
19	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuos/3]
20	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuos/4]
21	be/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[rcuos/5]
22	rt/4	root	0.00 B/s	0.00 B/s	0.00 %	0.00 %	[watchdog/0]



blktrace [DISK]

- Tool for generating/capturing disk I/O on block devices
- Select your block device, capture duration, output file

```
# blktrace -d /dev/md4 -w60 -o kvm_images
```

```
=== md4 ===
```

CPU 0:	286 events,	14 KiB data
CPU 1:	462 events,	22 KiB data
CPU 2:	471 events,	23 KiB data
CPU 3:	266 events,	13 KiB data
CPU 4:	44 events,	3 KiB data
CPU 5:	64 events,	4 KiB data
CPU 6:	91 events,	5 KiB data
CPU 7:	30 events,	2 KiB data
Total:	1714 events (dropped 0),	81 KiB data

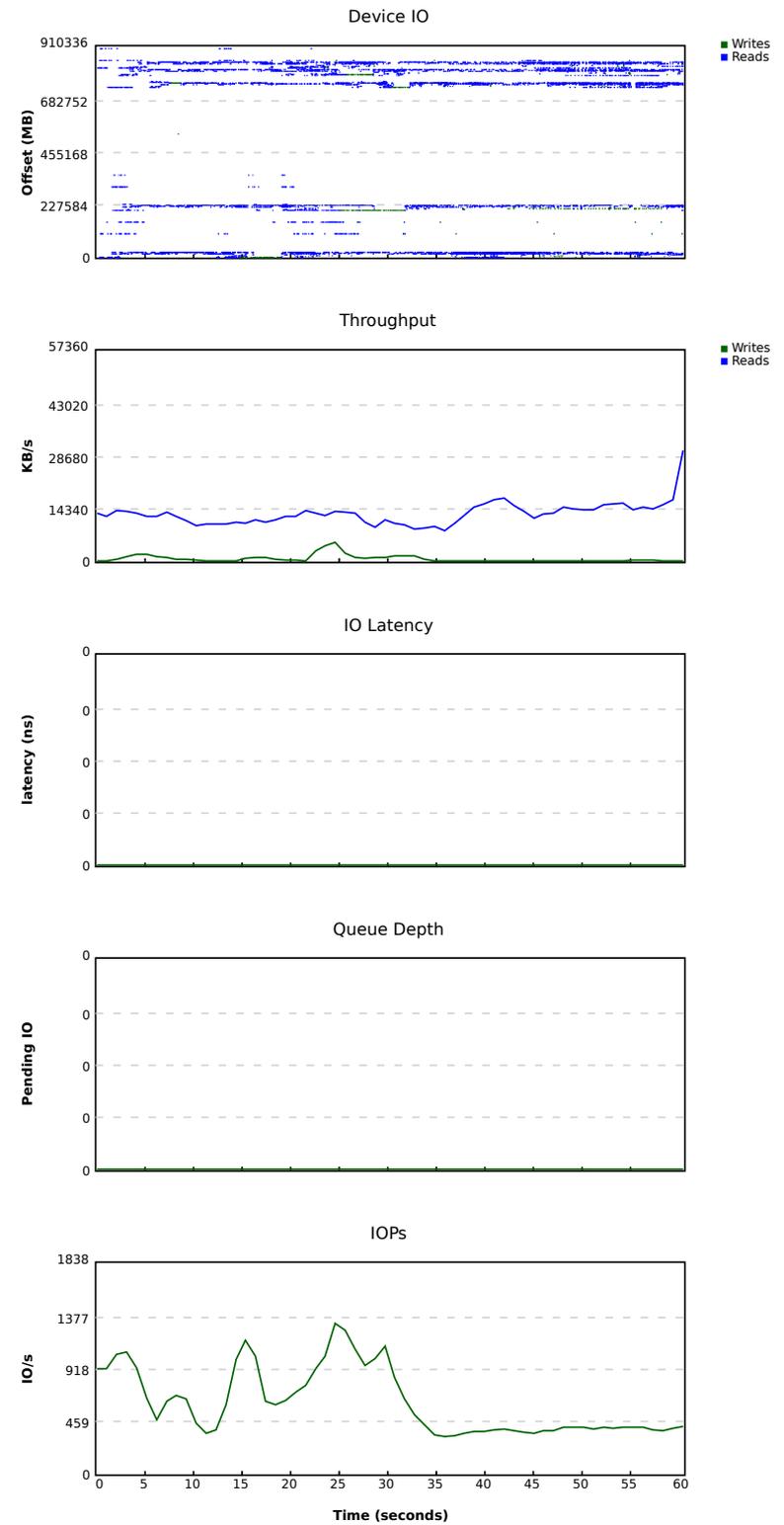


blktrace + iowatcher [DISK]

- Great, I have data but now what?
- blktrace captures data in binary format
- Need to read – graph?

```
# iowatcher -t kvm_images.blktrace -o  
kvm_images-trace.svg
```

- Generates a graph from captured data!
- Can even generate movie in ogg or mp4 format for visual playback of your data capture.



Cockpit [Multi-functional]

- Install, enable and start cockpit service on systems

```
# subscription-manager repos --enable rhel-7-server-extras-rpms
```

```
# yum install cockpit
```

```
# systemctl start cockpit
```

- Connect via port 9090 with web browser
- <https://myserver.com:9090>
- Enter local username/password



Cockpit :: Server View

- Connect to Cockpit agent via web browser on port 9090

RED HAT ENTERPRISE LINUX SERVER root

iscsi.i.skinnerla... Dashboard

System

- Services
- Logs
- Storage
- Networking
- Tools
- Subscriptions
- Accounts
- Diagnostic report
- Terminal

Hardware Gigabyte Technology Co.,...

Asset Tag To be filled by O.E.M.

Operating System CloudForms

Host Name iscsi.i.skinnerlabs.com

Domain

System Time 2016-06-06 14:52 ⓘ

Power Options ▾

Performance Profile throughput-performa...

% CPU

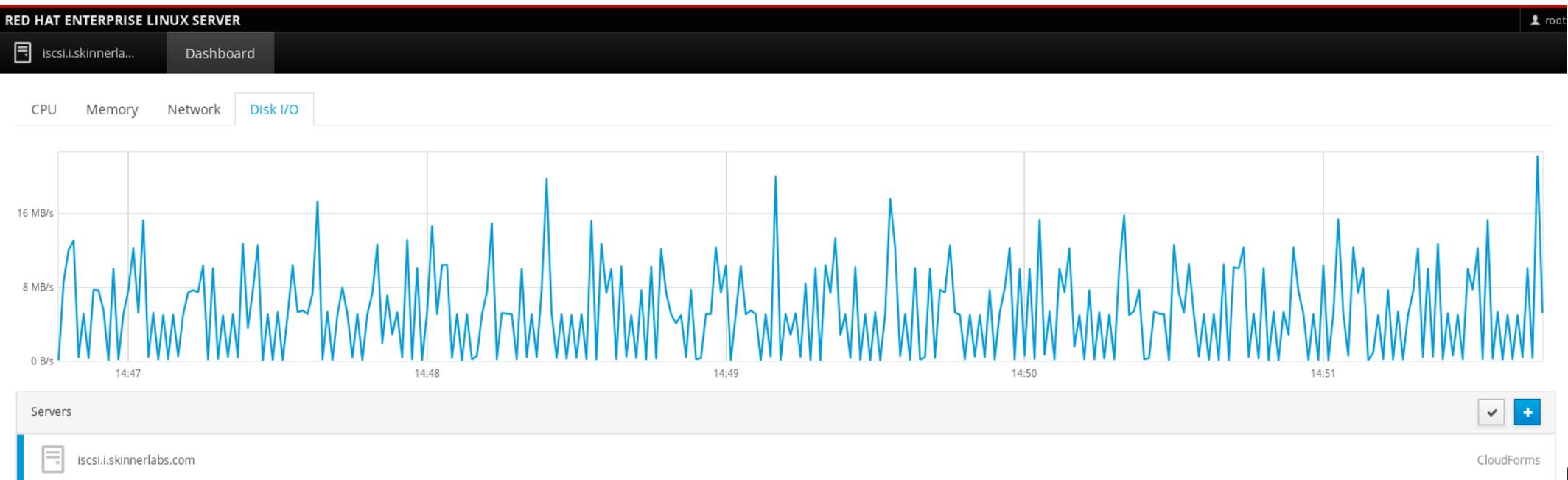
GB Memory

MB/s Disk I/O

Mbps Network Traffic

Cockpit :: Dashboard

- CPU :: Memory :: Network :: Disk I/O



PCP [Multi-functional]

- PCP = Performance Co-Pilot (RHEL 7 and RHEL => 6.6)
- RHEL 7 how to install:

```
# yum install pcp
```

```
# systemctl enable pmcd
```

```
# systemctl enable pmlogger
```

```
# systemctl start pmcd
```

```
# systemctl start pmlogger
```

- Data is collected every 60 seconds by default (sampled)
- To change, edit:
 - `/etc/pcp/pmlogger/control`
 - Append `-t 10s` to `LOCALHOSTNAME` line to change to 10 second interval
 - Restart pmlogger service



PCP :: Charts

- Extensive list of PCP metric chart add-ons

```
# yum search pcp | grep pmda
```

ActiveMQ
Apache
Bash shell
Bonded network
Cifs shell
Cisco shell
Device Mapper
389 Directory

Gfs2 shell
Gluster
GPFS Filesystem
Infiniband
JSON data
KVM
Lmsensors
Arbitrary log
Lustre
Mailq shell
Memcached
Mounts shell
MySQL

Named
Netfilter
Usenet News
NFS Clients
Nginx
Nvidia
Performance API
PowerDNS
Postfix
PostgreSQL
Roomtemp
Rpm shell
Rsyslog

Samba
Sendmail
Simple Network
Systemd
Trace shell
Unbound DNS
VMware
Weblog shell
Zimbra



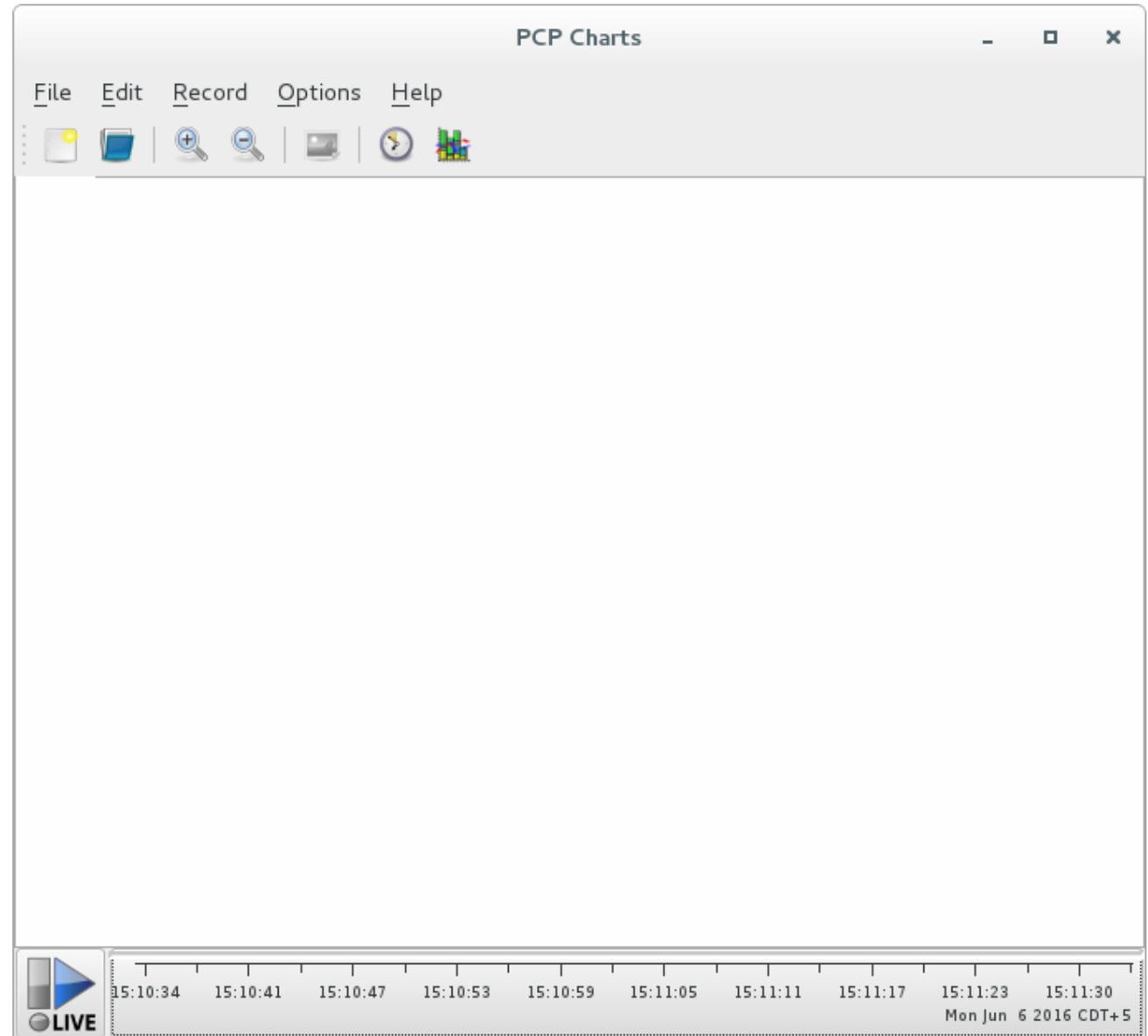
PCP :: Client/Server

- Client: pmchart
- Server: pmcd and pmlogger



PCP :: Client

- GUI client interface



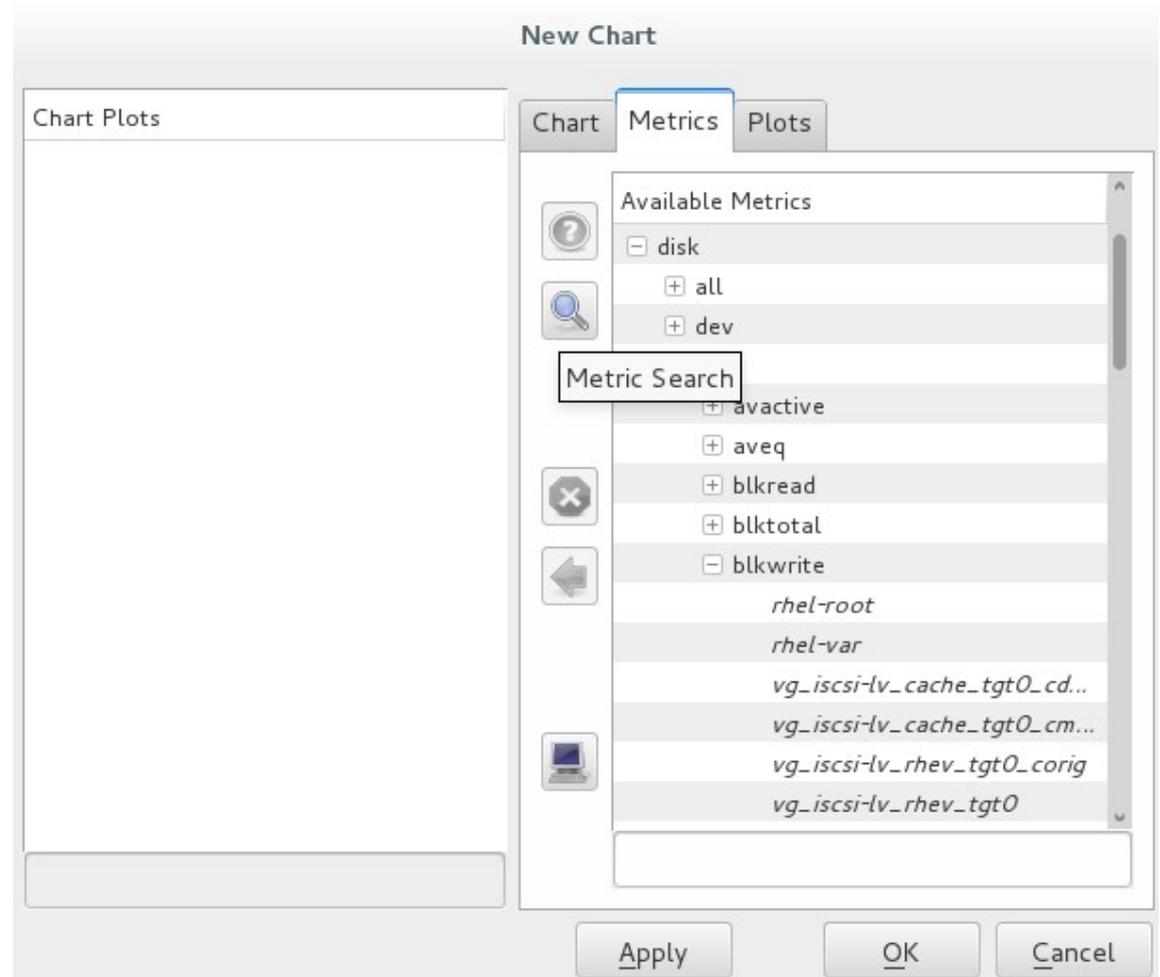
PCP :: Client

- Open pre-configured charts
- Select from Open View list



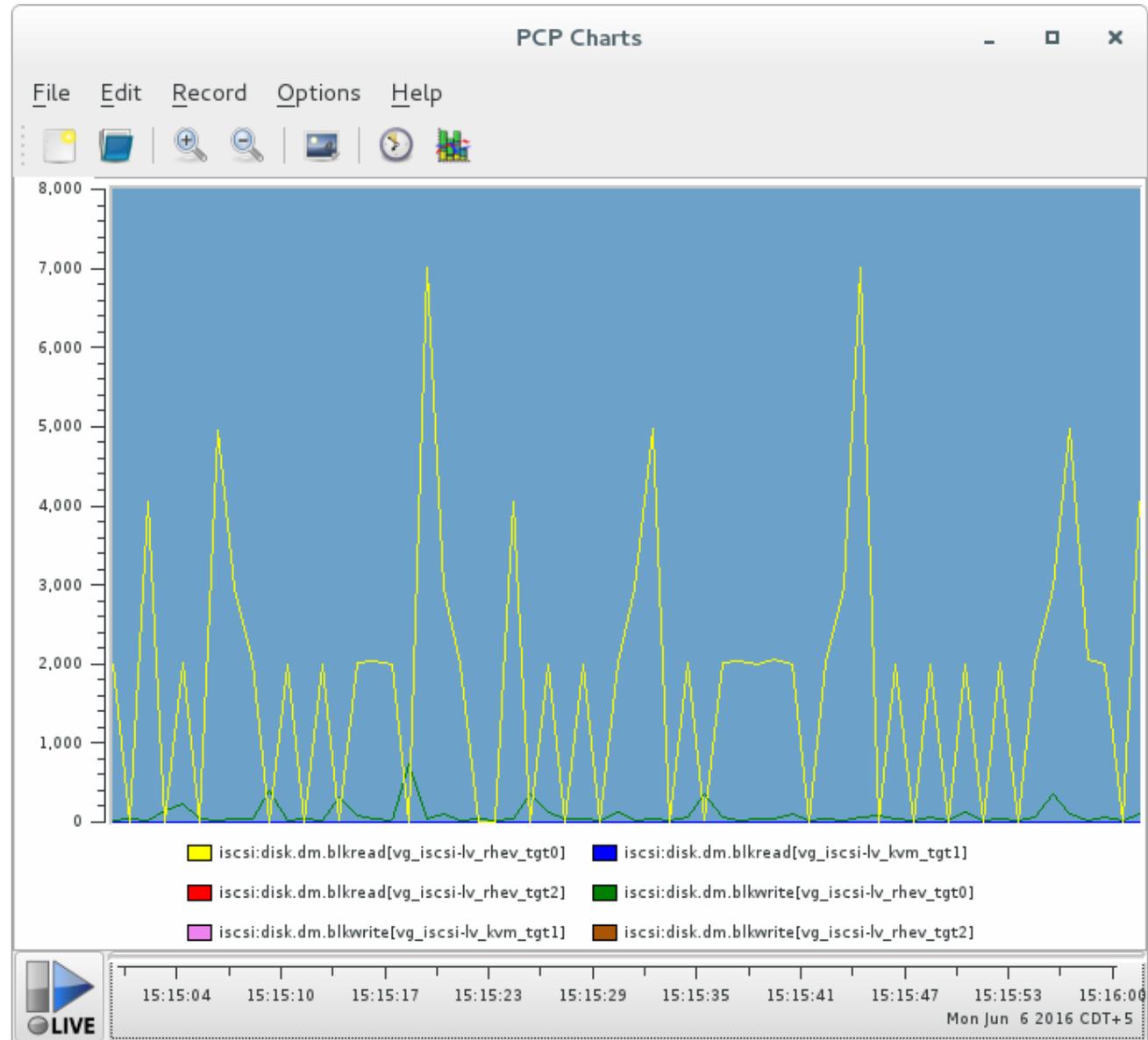
PCP :: Client

- Create new charts
- Choose host to monitor
- Select from available metrics
- Select 1 to n metric



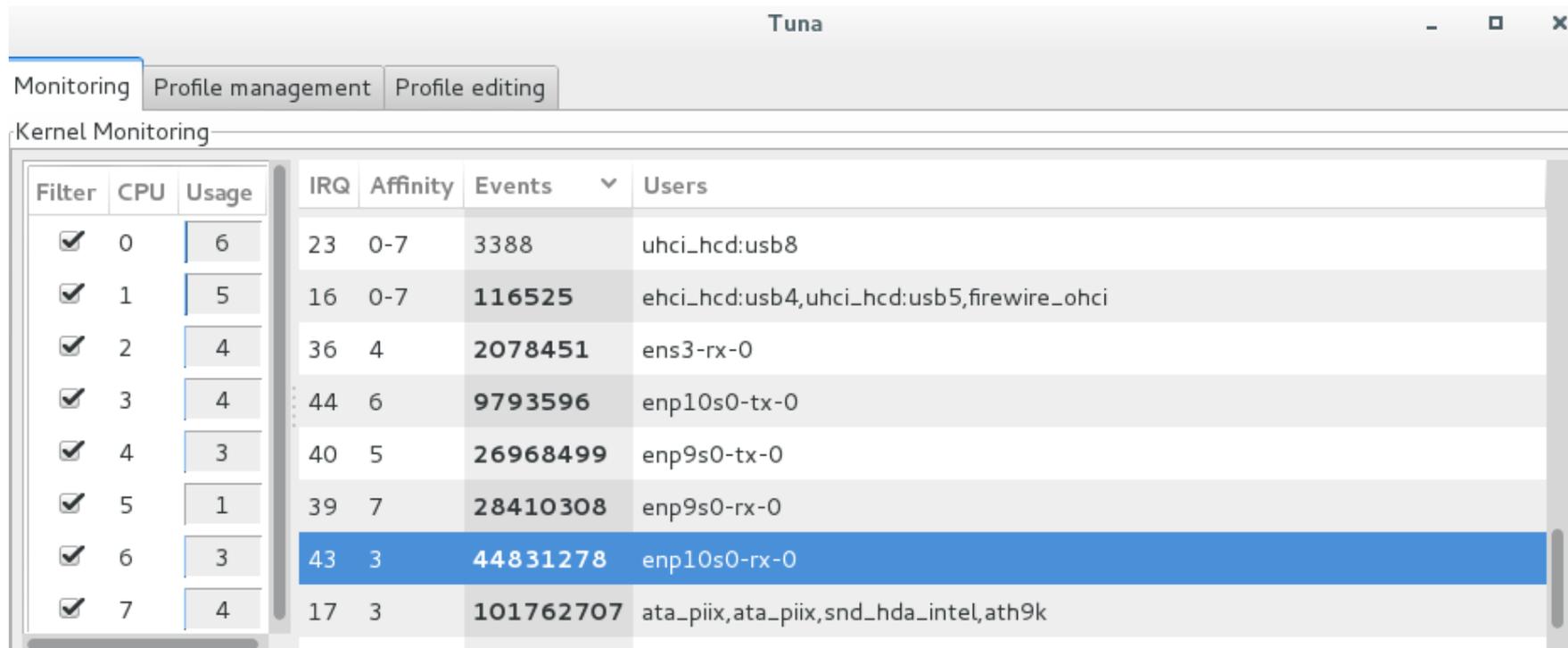
PCP :: Client

- Disk chart created
- Disk blkread
- Disk blkwrite
- Export as graphic
- Record for playback
- Import data from Collectd
- Export data to Webapps
 - Grafana
 - Graphite
 - Vector



tuna [Multi-functional]

- Monitor CPU/IRQ affinity
- Make changes, tweak
- RX/TX pinned to unique cores/threads



The screenshot shows the Tuna application window with three tabs: Monitoring, Profile management, and Profile editing. The Monitoring tab is active, displaying a table of kernel monitoring data. The table has columns for Filter, CPU, Usage, IRQ, Affinity, Events, and Users. The data is as follows:

Filter	CPU	Usage	IRQ	Affinity	Events	Users
<input checked="" type="checkbox"/>	0	6	23	0-7	3388	uhci_hcd:usb8
<input checked="" type="checkbox"/>	1	5	16	0-7	116525	ehci_hcd:usb4,uhci_hcd:usb5,firewire_ohci
<input checked="" type="checkbox"/>	2	4	36	4	2078451	ens3-rx-0
<input checked="" type="checkbox"/>	3	4	44	6	9793596	enp10s0-tx-0
<input checked="" type="checkbox"/>	4	3	40	5	26968499	enp9s0-tx-0
<input checked="" type="checkbox"/>	5	1	39	7	28410308	enp9s0-rx-0
<input checked="" type="checkbox"/>	6	3	43	3	44831278	enp10s0-rx-0
<input checked="" type="checkbox"/>	7	4	17	3	101762707	ata_piix,ata_piix,snd_hda_intel,ath9k



tuna [Multi-functional]

- Create a profile
- Change in real time
- Be careful – loaded gun!
- Save for later

The screenshot shows the Tuna application window with the following components:

- Monitoring** | **Profile management** | **Profile editing**
- Current active tuna profile: `example.conf`
- Buttons: **Save Snapshot**, **Save & Apply permanently**, **Restore changes**, **Apply changes**
- Kernel scheduler**
 - `kernel.core_pattern`: `||usr/libexec/abrt-hook-ccpp %s %c %p %`
 - `kernel.sched_autogroup_enabled`: 0
 - `kernel.sched_latency_ns`: 24000000
 - `kernel.sched_migration_cost_ns`: 5000000
 - `kernel.sched_min_granularity_ns`: 10000000
 - `kernel.sched_nr_migrate`: 32
 - `kernel.sched_rt_period_us`: 1000000
 - `kernel.sched_rt_runtime_us`: 950000
 - `kernel.sched_tunable_scaling`: 1
 - `kernel.sched_wakeup_granularity_ns`: 15000000
 - `kernel.sem`: 250 32000 32 128
- VM**
 - `vm.dirty_background_ratio`: 5
 - `vm.dirty_expire_centisecs`: 3000
 - `vm.dirty_ratio`: 40
 - `vm.dirty_writeback_centisecs`: 500
 - `vm.laptop_mode`: 0
 - `vm.max_map_count`: 65530
 - `vm.memory_failure_early_kill`: 0
 - `vm.swappiness`: 10
- Network IPv4**
 - `ipv4.conf.all.forwarding`: 0
 - `ipv4.conf.all.rp_filter`: 1
- Network IPv6**
 - `ipv6.conf.all.forwarding`: 0
 - `ipv6.conf.default.forwarding`: 0
 - `ipv6.conf.enp10s0.forwarding`: 0
 - `ipv6.conf.enp9s0.forwarding`: 0
 - `ipv6.conf.ens3.forwarding`: 0



Questions?

What are you using?