below

Interactive resource monitor for modern linux systems

Daniel Xu Facebook

Overview

- Background
- Features
- Tour of the UI
- Comparison
- Demo

Background

- Resource/system monitor
- Designed to replace atop (<u>https://www.atoptool.nl/</u>)
 - No cgroup awareness
 - May suffer priority inversions while host is under resource contention
 - On-disk data has custom delta compression and is easily corrupted
 - Unintuitive UX
- Developed by and for resource control team @ FB
 - Designed for post-mortem investigations
 - With input from major production atop users, cgroup/mm kernel developers, and container runtime developers
 - Completeness, reliability, and ease of use are core goals

Features

- Information regarding hardware resource utilization
- Viewing the cgroup hierarchy
 - Shortcuts for mapping cgroup -> processes
 - Pressure Stall Information (PSI)

Modes

- record mode to record system data (daemon mode)
 - Records _all_ processes in interval period, even those that have exited (using eBPF)
- replay mode to replay historical system data (client mode)
- live mode to view live system data (daemon + client mode)
- dump subcommand to report script-friendly information (eg json and csv) (client mode)

Tour of the UI

| | | Memory | CPU Pressure | Mem Pressure | I/O Pressure | Reads | Writ |
|---|--|--|--|--------------------------|--|---|---|
| | 202 400 | | 7.000 | | | | |
| | 161.44% | 13.8 GB | 7.23% | 0.00% | 0.00% | ? | ? |
| | | | | | | | ? |
| | | | | | | 2 | 7 |
| | | | | | | 2 | 7 |
| -runtime-dir@1001 service | | | | | | | ? |
| | | | | | | | ? |
| | | | | | | | ? |
| | | | | | | ż | 7 |
| | | | | | | ż | ? |
| | 0.00% | 816 KB | 0.00% | 0.00% | 0.00% | ? | ? |
| | 0.00% | 90.9 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| | 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| | 2clc6 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| ainer | 1.48% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| i con a con | e 001.slice r@1001.service sion-3.scope r-runtime-dir@1001.service ice cal.slice d-journald.service roker.service ervice e lice -8bda5b7cb0c34022cec6cbb44b6107552 | e 176.00% 176. | e 176.00% 9.5 GB 176.00% 9.4 GB 176.00% 9.4 GB 176.00% 9.4 GB sion-3.scope 0.00% 6 MB r-runtime-dir@1001.service 0.00% 0.0 B ice 0.49% 4.2 GB 105.7 MB 105.7 | 176.00% 9.5 GB 6.90% | e 176.00% 9.5 GB 6.90% 0.00% 000% 176.00% 9.4 GB 6.95% 0.00% 0.00% 176.00% 9.4 GB 6.95% 0.00% 0. | e 176.00% 9.5 GB 6.90% 0.00% 0.00% 0.00% 001.slice 176.00% 9.4 GB 6.95% 0.00% | e 176.00% 9.5 GB 6.90% 0.00% 0.00% ? 001.slice 176.00% 9.4 GB 6.95% 0.00% 0.00% ? 0176.00% 9.4 GB 6.95% 0.00% 0.00% ? 0176.00% 9.4 GB 6.95% 0.00% 0.00% ? 0176.00% 0.00% 0.00% 0.00% ? 0176.00% 0.00% 0.00% 0.00% ? 0176.00% 0.00% 0.00% 0.00% ? 0176.00% 0.00% 0.00% 0.00% ? 0176.00% 0.00% 0.00% 0.00% ? 0186.00% 0.00% 0.00% 0.00% ? 0187.00% 0.00% 0.00% 0.00% ? 0188.00% 0.00% 0.00% 0.00% ? 0198.00% 0.00% 0.00% 0.00% ? 0198.00% 0.00% 0.00% 0.00% ? 0198.00% 0.00% 0.00% 0.00% ? 0198.00% 0.00% 0.00% 0.00% ? 0198.00% 0.00% 0.00% 0.00% ? 0198.00% 0.00% 0.00% 0.00% ? 0199.00% 0.00% 0.00% 0.00% ? 0199.00% 0.00% 0.00% 0.00% ? 0199.00% 0.00% 0.00% 0.00% ? 0199.00% 0.00% 0.00% 0.00% ? |

Status bar

| PU | Usage | 42.14% | | User | 23.24% | | System | 18.49% | | |
|---------------|---------|---------|-----------|----------|----------|-----------|---------|---------|----------|---------|
| 1em | Total | 15.4 GB | | Free | 1.1 GB | | Anon | 9.4 GB | File | 3.9 GB |
| /M | Page In | 0.0 B/s | | Page Out | 3.2 MB/s | | Swap In | 0.0 B/s | Swap Out | 0.0 B/s |
| I/O (Rd Wr) | dm-0 | 0.0 B/s | 819.2 B/s | nvme0n1 | 0.0 B/s | 819.2 B/s | | | | |
| Iface (Rx Tx) | eth0 | 35.6 B | 0.0 B | lo | 0.0 B | 0.0 B | | | | |

0.1.0

replay

8bda5b7cb0c3

05/14/2021 16:32:16 UTC+00:00

Elapsed: 5s

| ame | CPU | Memory | CPU Pressure | Mem Pressure | I/O Pressure | Reads | Write |
|--|---------|----------|--------------|--------------|--------------|-------|-------|
| root> | 161.44% | 13.8 GB | 7.23% | 0.00% | 0.00% | ? | ? |
| - user.slice | 176.00% | 9.5 GB | 6.90% | 0.00% | 0.00% | ? | ? |
| └ user-1001.slice | 176.00% | 9.4 GB | 6.95% | 0.00% | 0.00% | ? | ? |
| L+ user@1001.service | 176.01% | 9.4 GB | 6.94% | 0.00% | 0.00% | ? | ? |
| └ session-3.scope | 0.00% | 6 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ user-runtime-dir@1001.service | 0.00% | 0.0 B | 0.00% | 0.00% | 0.00% | ? | ? |
| + system.slice | 0.49% | 4.2 GB | 0.20% | 0.00% | 0.00% | ? | ? |
| - hostcritical.slice | 0.01% | 105.7 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └─ systemd-journald.service | 0.01% | 91.4 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └─ dbus-broker.service | 0.01% | 13.4 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ sshd.service | 0.00% | 816 KB | 0.00% | 0.00% | 0.00% | ? | ? |
| + init.scope | 0.00% | 90.9 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| – machine.slice | 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ libpod-8bda5b7cb0c34022cec6cbb44b6107552c1c6 | | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ container | 1.48% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| | | | | | | | |

| Usage 42.14% Total 15.4 GB Page In 0.0 B/s (Rd Wr) dm-0 0.0 B/s 819.2 B/s ce (Rx Tx) eth0 35.6 B 0.0 B | User 23.24% Free 1.1 GB Page Out 3.2 MB nvme0n1 0.0 B/ lo 0.0 B | /s | Anon 9.4 Swap In 0.0 | .49% 4 GB 9 B/s | File 3. Swap Out 0. | 9 GB 9 B/s | |
|--|---|--------------------|-------------------------|-----------------------|------------------------|---------------|------|
| eral CPU Mem I/O Pressure | CDU | | CRIL B | | T (0. B | | |
| e | CPU | Memory | CPU Pressure | Mem Pressure | I/O Pressure | Reads | Writ |
| ot> | 161.44% | 13.8 GB | 7.23% | 0.00% | 0.00% | ? | ? |
| user.slice | 176.00% | 9.5 GB | 6.90% | 0.00% | 0.00% | ? | ? |
| └ user-1001.slice | 176.00% | 9.4 GB | 6.95% | 0.00% | 0.00% | ? | ? |
| L+ user@1001.service | 176.01% | 9.4 GB | 6.94% | 0.00% | 0.00% | ? | ? |
| └ session-3.scope | 0.00% | 6 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ user-runtime-dir@1001.service | 0.00% | 0.0 B | 0.00% | 0.00% | 0.00% | ? | ? |
| system.slice | 0.49% | 4.2 GB | 0.20% | 0.00% | 0.00% | ? | ? |
| hostcritical.slice | 0.01% | 105.7 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ systemd-journald.service | 0.01% | 91.4 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ dbus-broker.service | 0.01% | 13.4 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └─ sshd.service | 0.00% | 816 KB | 0.00% | 0.00% | 0.00% | ? | ? |
| init.scope | 0.00% | 90.9 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| machine.slice | 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| libpod-8bda5b7cb0c34022cec6cbb44b61075 | 52c1c6 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ container | 1.48% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| machine.slice └─ libpod-8bda5b7cb0c34022cec6cbb44b61075 | 1.46% 52c1c6 1.46% | 19.1 MB 19.1 MB | 0.00% 0.00% | 0.00% 0.00% | 0.00% 0.00% | | |

System overview

```
Elapsed: 5s
05/14/2021 16:32:16 UTC+00:00
                                                                  8bda5b7cb0c3
                                                                                        0.1.0
                                                                                                      replay
                               42.14%
15.4 GB
                                                                                                   System 18.49%
Anon 9.4 GB
CPU
                    Usage
Total
                                                           User
                                                                       23.24%
                                                                       1.1 GB
                                                                                                                                           File
                                                                                                                                                      3.9 GB
                                                            Free
VM Page In 0.0 B/s
I/O (Rd|Wr) dm-0 0.0 B/s
Iface (Rx|Tx) eth0 35.6 B
                                             Page Out 3.2 MB/s
|819.2 B/s nvmeOn1 0.0 B/s
|0.0 B lo 0.0 B
                                                                                                   Swap In 0.0 B/s
                                                                                                                                           Swap Out 0.0 B/s
                                                                                    |819.2 B/s
|0.0 B
```

| Name | CPU | Memory | CPU Pressure | Mem Pressure | I/O Pressure | Reads | Write |
|---|---------|----------|--------------|--------------|--------------|-------|-------|
| <root></root> | 161.44% | 13.8 GB | 7.23% | 0.00% | 0.00% | ? | ? |
| └ user.slice | 176.00% | 9.5 GB | 6.90% | 0.00% | 0.00% | ? | ? |
| └ user-1001.slice | 176.00% | 9.4 GB | 6.95% | 0.00% | 0.00% | ? | ? |
| 4 user@1001.service | 176.01% | 9.4 GB | 6.94% | 0.00% | 0.00% | ? | ? |
| └ session-3.scope | 0.00% | 6 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ user-runtime-dir@1001.service | 0.00% | 0.0 B | 0.00% | 0.00% | 0.00% | ? | ? |
| L+ system.slice | 0.49% | 4.2 GB | 0.20% | 0.00% | 0.00% | ? | ? |
| └─ hostcritical.slice | 0.01% | 105.7 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ systemd-journald.service | 0.01% | 91.4 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ dbus-broker.service | 0.01% | 13.4 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └─ sshd.service | 0.00% | 816 KB | 0.00% | 0.00% | 0.00% | ? | ? |
| L+ init.scope | 0.00% | 90.9 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ machine.slice | 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └─ libpod-8bda5b7cb0c34022cec6cbb44b6107552c1c6 | 1.46% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |
| └ container | 1.48% | 19.1 MB | 0.00% | 0.00% | 0.00% | ? | ? |

cgroup view

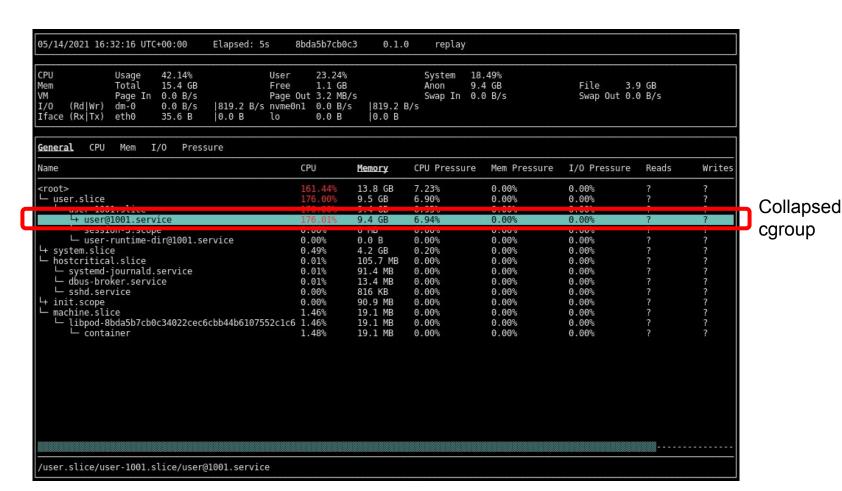
/user.slice/user-1001.slice/user@1001.service

| | | Swap Out 0.0 | 9 B/s | Anon 9.4 Swap In 0.0 3/s | | 1.1 GB Out 3.2 MB/ On1 0.0 B/s 0.0 B | Free Page (819.2 B/s nvme0 0.0 B lo | Total 15.4 GB Page In 0.0 B/s dm-0 0.0 B/s eth0 35.6 B | M /O (Rd Wr) face (Rx Tx) |
|---------|-------------|----------------------------------|----------------------------------|--------------------------------|-------------------------------|---|---|--|---|
| 11/21 | . 0 | T/0 B | M | CRIL B | | CDU | sure | Mem I/O Press | eneral CPU |
| s Write | e Reads | I/O Pressure | Mem Pressure | CPU Pressure | Memory | CPU | | | ame |
| ? | ? | 0.00% | 0.00% | 7.23% | 13.8 GB | 161.44% | | | root> |
| ? | ? | 0.00% | 0.00% | 6.90% | 9.5 GB | 176.00% | | | - user.slice |
| ? | ? | 0.00% | 0.00% | 6.95% | 9.4 GB | 176.00% | | | ∟ user-100 |
| ? | ? | 0.00% | 0.00% | 6.94% | 9.4 GB | 176.01% | | 1001.service | |
| ? | ? | 0.00% | 0.00% | 0.00% | 6 MB | 0.00% | | on-3.scope | |
| ? | ? | 0.00% | 0.00% | 0.00% | 0.0 B | 0.00% | ervice | runtime-dir@1001.se | ∟ user- |
| ? | ? | 0.00% | 0.00% | 0.20% | 4.2 GB | 0.49% | | e | <pre>+ system.slic</pre> |
| ? | ? | 0.00% | 0.00% | 0.00% | 105.7 MB | 0.01% | | | hostcritica |
| ? | ? | 0.00% | 0.00% | 0.00% | 91.4 MB | 0.01% | | journald.service | |
| ? | ? | 0.00% | 0.00% | 0.00% | 13.4 MB | 0.01% | | ker.service | └ dbus-bro |
| ? | ? | 0.00% | 0.00% | 0.00% | 816 KB | 0.00% | | vice | |
| ? | ? | 0.00% | 0.00% | 0.00% | 90.9 MB | 0.00% | | | init.scope |
| ? | ? | 0.00% | 0.00% | 0.00% | 19.1 MB | 1.46% | | | - machine.sli |
| ? | ? | 0.00% | 0.00% | 0.00% | 19.1 MB | 1.46% | 6cbb44b6107552c1c6 | bda5b7cb0c34022cec6 | └─ libpod-8 |
| ? | ? | 0.00% | 0.00% | 0.00% | 19.1 MB | 1.48% | | iner | └ conta |
| | ? ? ? | 0.00% 0.00% 0.00% 0.00% | 0.00% 0.00% 0.00% 0.00% | 0.00% 0.00% 0.00% | 90.9 MB 19.1 MB 19.1 MB | 0.00% 0.00% 1.46% 5 1.46% | 5cbb44b6107552clc6 | ce bda5b7cb0c34022cec6 | └─ sshd.ser + init.scope - machine.sli └─ libpod-8 |

Resource tabs

| CPU Usage 42.14% User Mem Total 15.4 GB Free | 23.24% 1.1 GB Out 3.2 MB, n1 0.0 B/ 0.0 B | /s | System 18 Anon 9.4 Swap In 0.0 | .49% 4 GB 0 B/s | File 3. Swap Out 0. | 9 GB 0 B/s | | |
|---|---|--|---|--|---|---------------|---|-------------|
| General CPU Mem I/O Pressure | | | | | | | | |
| Name | CPU | Memory | CPU Pressure | Mem Pressure | I/O Pressure | Reads | Writes | Data in tab |
| <pre><root> user.slice user.l001.slice usereql0001.service session-3.scope user-runtime-dir@1001.service system.slice hostcritical.slice systemd-journald.service dbus-broker.service sshd.service init.scope machine.slice libpod-8bda5b7cb0c34022cec6cbb44b6107552c1c6 container</root></pre> | 161.44% 176.00% 176.00% 176.01% 0.00% 0.00% 0.01% 0.01% 0.01% 0.01% 0.01% 1.46% 1.46% | 13.8 GB 9.5 GB 9.4 GB 9.4 GB 6 MB 0.0 B 4.2 GB 105.7 MB 91.4 MB 13.4 MB 816 KB 90.9 MB 19.1 MB 19.1 MB 19.1 MB | 7.23% 6.90% 6.95% 6.94% 0.00% 0.20% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% | 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% | 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% | | ? | |
| | | | | | | | | Scrollable |

/user.slice/user-1001.slice/user@1001.service



Expanded cgroup

| 5/14/2021 16: | 32:16 UTC | +00:00 | Elapsed: 5s | 8bc | la5b7cb0c3 | 0.1.0 | repla | у | | | | |
|---------------|-----------|---------|-------------|------------|------------|-------------|---------|---------|---------|----------|---------|--------|
| PU | Usage | 42.14% | | User | 23.24% | | System | 18.49% | | | | |
| em | Total | 15.4 GB | | Free | 1.1 GB | | Anon | 9.4 GB | | File | 3.9 GB | |
| | Page In | 0.0 B/s | | Page Out | 3.2 MB/s | | Swap In | 0.0 B/s | | Swap Out | 0.0 B/s | |
| 0 (Rd Wr) | dm-0 | 0.0 B/s | | nvme0n1 | | 819.2 B/s | | | | | | |
| ace (Rx Tx) | eth0 | 35.6 B | 0.0 B | lo | 0.0 B | 0.0 B | | | | | | |
| neral CPU | Mem I | /0 | | | | | | | | | | |
| mm | | (| Cgroup | | | | | Pid | Ppid | State | CPU | CPU Us |
| b Content | | | /e/qnome-la | unched-f | irefox.des | ktop-204660 | 3.scope | 2048151 | 2046603 | SLEEPING | 4.40% | 3.40% |
| refox | | | /e/gnome-la | unched - f | irefor dec | kton-204660 | 3 scone | 2046603 | 3631 | SLEEPING | 64.40% | 30.20% |

Process view

|packagekitd |element-desktop SLEEPING /...l.service/app-flatpak-im.riot.Riot-54267.scope 54331 54283 SLEEPING 0.80% 0.60% /...e/gnome-launched-firefox.desktop-2046603.scope 2048129 2046603 SLEEPING 2.80% 2.80% Web Content WebExtensions /...e/gnome-launched-firefox.desktop-2046603.scope 2046806 2046603 SLEEPING 0.20% 0.20% /...e/gnome-launched-firefox.desktop-2046603.scope 2048242 Web Content 2046603 SLEEPING 0.00% 0.00% /...e/gnome-launched-firefox.desktop-2046603.scope 2046781 2046603 SLEEPING Web Content 0.20% 0.20% /...e/gnome-launched-firefox.desktop-2046603.scope 2048300 Web Content 2046603 RUNNING 98.40% 58.20% /...vice/gnome-launched-dropbox.desktop-4266.scope 3728781 SLEEPING dropbox 3307 0.20% 0.20% /...e/gnome-launched-firefox.desktop-2046603.scope 2057416 SLEEPING Web Content 2046603 0.20% 0.00% /...e/gnome-launched-firefox.desktop-2046603.scope 2644171 2046603 SLEEPING Web Content 0.00% 0.00% gnome-software /...ched-gnome-software-service.desktop-4252.scope 4252 3596 SLEEPING 0.00% 0.00% /.../user@1001.service/gnome-shell-wayland.service 3631 anome-shell 3307 SLEEPING 0.00% 0.00% /...e/gnome-launched-firefox.desktop-2046603.scope 2787743 Web Content 2046603 SLEEPING 0.00% 0.00% systemd-journal /hostcritical.slice/systemd-journald.service SLEEPING 0.00% 0.00% Privileged Cont /...e/gnome-launched-firefox.desktop-2046603.scope 2046970 2046603 SLEEPING 0.00% 0.00% systemd 0 SLEEPING 0.00% /init.scope 0.00% /...d-home-dlxu-.cargo-bin-alacritty-4131527.scope 4131527 SLEEPING alacritty 3307 0.00% 0.00% gnome-control-c /...ched-gnome-network-panel.desktop-2954586.scope 2954586 3631 SLEEPING 0.00% 0.00% /system.slice/below.service 4137616 SLEEPING 0.00% 0.00% podman gsd-media-keys /...slice/user@1001.service/gsd-media-keys.service 4129 3307 SLEEPING 0.00% 0.00% element-desktop /...l.service/app-flatpak-im.riot.Riot-54148.scope 54250 54249 SLEEPING 0.20% 0.20% element-desktop /...l.service/app-flatpak-im.riot.Riot-54148.scope 54310 54263 SLEEPING 0.00% /usr/lib64/firefox/firefox -contentproc -childID 9 -isForBrowser -prefsLen 5839 -prefMapSize 252334 -parentBuildID 20210422145831 -appo

/usr/lib64/firefox/browser 2046603 true tab

Comparison

| | Terminal interface | Host level stats | cgroup support | Record/replay | Scriptable | Web interface |
|----------|--------------------|------------------|----------------|---------------|------------|---------------|
| atop | Yes | Yes | No | Yes | Yes | No |
| cAdvisor | No | No | cgroup1 only | Yes | Yes | Yes |
| htop | Yes | Yes | No | No | No | No |
| below | Yes | Yes | cgroup2 only | Yes | Yes | No |

By no means complete!

Demo

Thanks bye

Repository:

https://github.com/facebookincubator/resctl