

Cool "fičury" v systemd

Václav Pavlín, Lukáš Nykrýn

Journal

“Nový přístup k logování, který poskytuje strukturované, indexované a spolehlivé zaznamenávání událostí v systémech využívajících systemd.”

systemctl status

```
vpavlin@unused-4-132:~  
root@localhost:~  
root@localhost:~  
vpavlin@unused-4-132:~ 132x32  
[vpavlin@unused-4-132 ~]  
$ systemctl status systemd-logind.service  
systemd-logind.service - Login Service  
  Loaded: loaded (/usr/lib/systemd/system/systemd-logind.service; static)  
  Active: active (running) since So 2013-11-02 15:45:40 CET; 16h ago  
    Docs: man:systemd-logind.service(8)  
          man:logind.conf(5)  
          http://www.freedesktop.org/wiki/Software/systemd/multiseat  
Main PID: 558 (systemd-logind)  
  CGroup: name=systemd:/system/systemd-logind.service  
          └─558 /usr/lib/systemd/systemd-logind  
  
lis 02 15:52:00 unused-4-132.brq.redhat.com systemd-logind[558]: New session 2 of user root.  
lis 02 15:52:00 unused-4-132.brq.redhat.com systemd-logind[558]: Removed session 2.  
lis 02 21:45:44 unused-4-132.brq.redhat.com systemd-logind[558]: Suspend key pressed.  
lis 02 21:45:46 unused-4-132.brq.redhat.com systemd-logind[558]: Lid closed.  
lis 03 05:10:32 unused-4-132.brq.redhat.com systemd-logind[558]: Lid opened.  
lis 03 05:10:32 unused-4-132.brq.redhat.com systemd-logind[558]: Operation finished.  
lis 03 07:25:40 unused-4-132.brq.redhat.com systemd-logind[558]: Suspend key pressed.  
lis 03 07:56:23 unused-4-132.brq.redhat.com systemd-logind[558]: Lid closed.  
lis 03 07:56:23 unused-4-132.brq.redhat.com systemd-logind[558]: Lid opened.  
lis 03 07:56:23 unused-4-132.brq.redhat.com systemd-logind[558]: Operation finished.  
[vpavlin@unused-4-132 ~]  
$
```

Journal

- text vs. strukturovaná data
- podvrhnutelnost vs. bezpečnost
- rotace podle času vs. podle využití místa
- indexace, integrace...
- “od bootu až do vypnutí”

Journal

- v paměti, nebo na disku
- pole journalu
- API - C, Python, PHP, Node.js

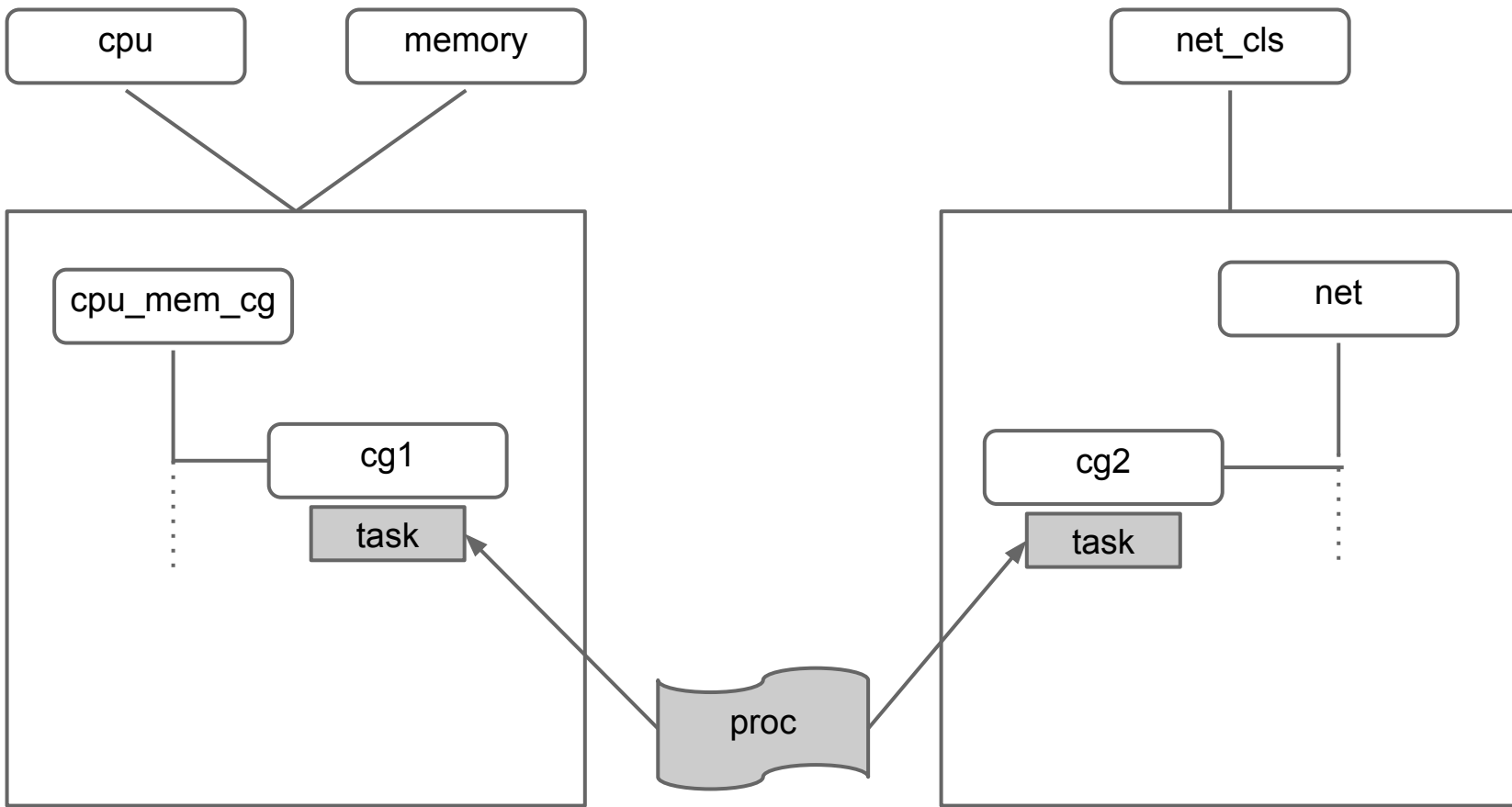
```
int sd_journal_print (int priority, const char* format,...);  
int sd_journal_send (const char* format,...);
```

Journal fields

`_SERVICE=systemd-logind.service`
`MESSAGE=User harald logged in`
`MESSAGE_ID=422bc3d271414bc8bc9570f222f24a9`
`_EXE=/lib/systemd/systemd-logind`
`_COMM=systemd-logind`
`_CMDLINE=/lib/systemd/systemd-logind`
`_PID=4711`
`_UID=0`
`_GID=0`
`_SYSTEMD_CGROUP=/system/systemd-logind.service`
`_CGROUPS=cpu:/system/systemd-logind.service`
`PRIORITY=6`
`_BOOT_ID=422bc3d271414bc8bc95870f222f24a9`
`_MACHINE_ID=c686f3b205dd48e0b43ceb6eda479721`
`_HOSTNAME=waldi`
`LOGIN_USER=500`

Control Groups

- Mechanismus jádra pro slučování / rozdělování procesů a jejich potomků do skupin
- Task
 - proces
- Group
 - skupina procesů
- Hierachy
 - strom skupin
- Controllers



Mount + přímý zápis

```
mount -t tmpfs cgroup_root /sys/fs/cgroup
mkdir /sys/fs/cgroup/cpuset
mount -t cgroup cpuset -ocpuset /sys/fs/cgroup/cpuset
cd /sys/fs/cgroup/cpuset
mkdir Charlie
cd Charlie
/bin/echo 2-3 > cpuset.cpus
/bin/echo 1 > cpuset.mems
/bin/echo $$ > tasks
```

cgconfig

```
group daemons/sql {  
    perm {  
        task {  
            uid = root;  
            gid = sqladmin;  
        } admin {  
            uid = root;  
            gid = root;  
        }  
    }  
    } cpu {  
        cpuset.mems = 0;  
        cpuset.cpus = 0;  
    }  
}
```

All your cgroups are belong to us!

-- Lennart Poettering

- pouze jeden strom
- pouze jeden zapisovatel z user-space

systemd's Resource Control Concepts

- **Service**
 - služby spuštěné systemd
- **Slice**
 - seskupování service a scope do stromové hierarchie
- **Scope**
 - vytvořené pouze přes api
 - nevytváří vlastní procesy
 - externě vytvořené procesy se do scope “registrují” přes api

Service

[Unit]

Description=My little test daemon

[Service]

Type=simple

ExecStart=/usr/bin/my_little_daemon

Slice=system-delicious.slice

Slice

[Unit]

Description=My slice

Before=slices.target

[Slice]

Slice=system.slice

MemoryLimit=1M

CPUShares=512

```
└─user.slice
  └─user-0.slice
    ├──session-1.scope
    │   ├──24 login -- root
    │   └─28 -bash
    └─user@0.service
        ├──26 /usr/lib/systemd/systemd --user
        └─27 (sd-pam)
└─system.slice
  ├──1 /usr/lib/systemd/systemd
  ├──system-delicious.slice
  │   ├──my_little_daemon.service
  │   │   └─48 /usr/bin/my_little_daemon
  ├──systemd-logind.service
  │   └─23 /usr/lib/systemd/systemd-logind
  ├──dbus.service
  │   └─22 /bin/dbus-daemon --system --address=systemd:
  ├──systemd-journald.service
  │   └─12 /usr/lib/systemd/systemd-journald
```

systemd-run

```
# systemd-run env
```

```
Running as unit run-19945.service.
```

```
# journalctl -u run-19945.service
```

```
Sep 08 07:37:21 bupkis systemd[1]: Starting /usr/bin/env...
```

```
Sep 08 07:37:21 bupkis systemd[1]: Started /usr/bin/env.
```

```
Sep 08 07:37:21 bupkis env[19948]: PATH=/usr/local/sbin:  
/usr/local/bin:/usr/sbin:/usr/bin
```

```
Sep 08 07:37:21 bupkis env[19948]: LANG=en_US.UTF-8
```

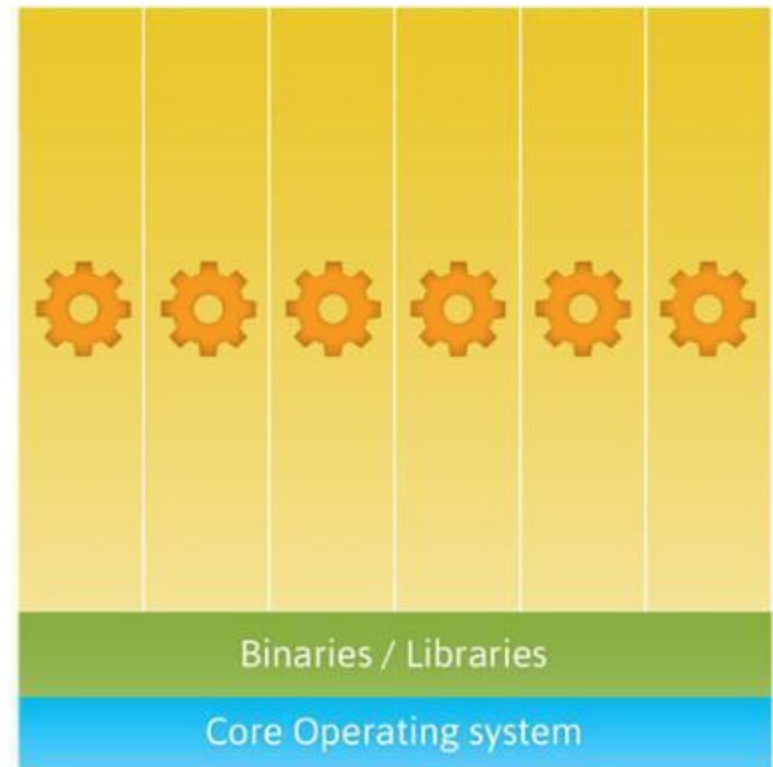
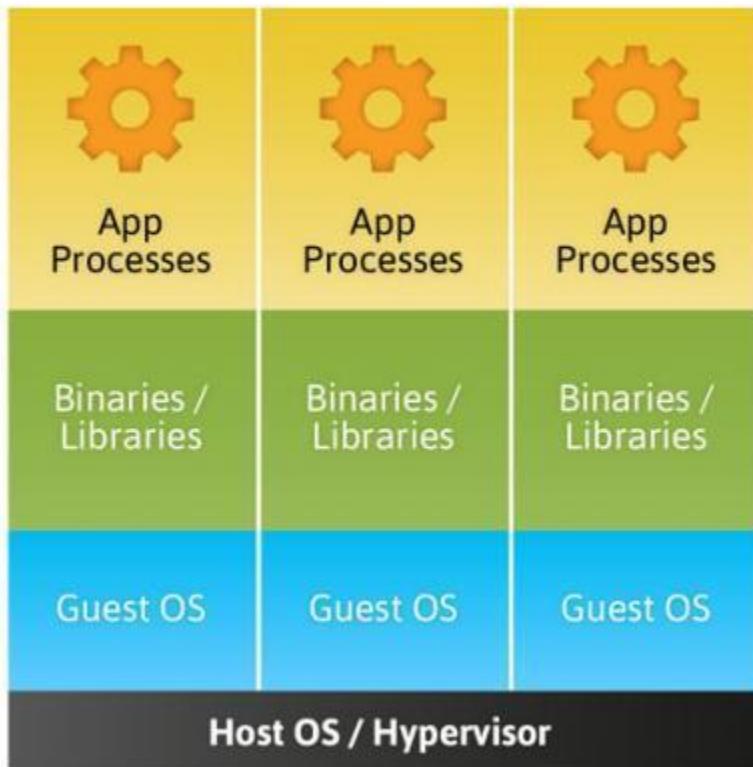
```
Sep 08 07:37:21 bupkis env[19948]:
```

```
BOOT_IMAGE=/vmlinuz-3.11.0-0.rc5.git6.2.fc20.x86_64
```


Kontejnery

“systemd-nspawn je chroot na steroidech”

VM vs. kontejnery



Kontejnery

Hodí se pro

- vývoj
- testování
- startupy (Pantheon)
- ...

Zdroje

- <http://0pointer.de/blog/projects/journalctl.html>
- <https://www.kernel.org/doc/Documentation/cgroups/cgroups.txt>
- <http://www.freedesktop.org/software/systemd/man/systemd.resource-control.html>
- <http://www.freedesktop.org/wiki/Software/systemd/ControlGroupInterface/>
- <http://www.linuxjournal.com/content/containers%E2%80%94not-virtual-machines%E2%80%94are-future-cloud>
- <http://0pointer.de/public/systemd-man/systemd-nspawn.html>

Dotazy?