Subject: cgroup release_agent issues with OpenVZ kernel
Posted by potatosalad on Tue, 01 May 2012 17:55:37 GMT
View Forum Message <> Reply to Message

I've filed a bug report at bugzilla.openvz.org/show_bug.cgi?id=2261

Here's my setup on a CentOS 6 host with...

kernel: 2.6.32-042stab053.5
vzctl: vzctl-3.1-87.git.5fa6a80.x86_64
vzctl-lib: vzctl-lib-3.1-87.git.5fa6a80.x86_64
ploop: ploop-1.1-1.x86_64
ploop-lib: ploop-lib-1.1-1.x86_64

I'm unable to get the notify_on_release and release_agent features of cgroups to work when using the OpenVZ kernel. (see sections 1.4 and 2.1 of the cgroup documentation)

The following works perfectly when using the regular CentOS 6 kernel (2.6.32-220.el6.x86_64)...

Step 1: create the release_agent script

```
[root@centos]# cat > /root/release-agent <<-'END'
#!/bin/sh

echo "$(date) - $@" | tee -a /tmp/my-cgroup.log

exit 0

END
[root@centos]# chmod +x /root/release-agent
[root@centos]# /root/release-agent testing
[root@centos]# cat /tmp/my-cgroup.log
Tue May  1 11:27:08 MDT 2012 - testing
```

Step 2: mount a named cgroup without any subsystems loaded (just used for labeling/grouping processes)

```
[root@centos]# mkdir /my-cgroup
[root@centos]# mount -t cgroup -o none,name=my-cgroup cgroup /my-cgroup
[root@centos]# /bin/echo 1 > /my-cgroup/notify_on_release
[root@centos]# /bin/echo /root/release-agent > /my-cgroup/release_agent
```

Step 3: create a script that loops so we can throw it in a child cgroup

```
[root@centos]# cat > /root/my-loop <<-'END'
#!/bin/bash

CONTINUE=1

trap "CONTINUE=0" SIGINT SIGTERM

echo "starting"

while [ $CONTINUE = 1 ]
do
  echo $$
  sleep 1
done

echo "stopping"

exit 0

END
[root@centos]# chmod +x /root/my-loop
```

Step 4: run the loop script and add the pid to a cgroup named "my-loop"

```
[root@centos]# /root/my-loop 2>&1 > /tmp/my-loop.log &
[1] 1507
[root@centos]# mkdir /my-cgroup/my-loop
[root@centos]# /bin/echo 1507 > /my-cgroup/my-loop/tasks
[root@centos]# cat /my-cgroup/my-loop/tasks
1507
1581
```

Step 5: send TERM to the main pid of the loop, which will empty the cgroup, firing the release_agent

```
[root@centos]# kill -s TERM 1507
[root@centos]# cat /tmp/my-cgroup.log
Tue May  1 11:27:08 MDT 2012 - testing
Tue May  1 11:29:34 MDT 2012 - /my-loop
```

When using the OpenVZ kernel (2.6.32-042stab053.5) everything works great until Step 5, the release_agent is never actually fired.

Note: when using the fedora-16 template with systemd, I am able to get release_agent to fire correctly only inside the container's /sys/fs/cgroup/systemd, but not inside any other cgroups on the host or within the containers. Is there something that OpenVZ does to allow systemd to do this, but disallow it elsewhere?

Has anyone had any experience with getting release_agent and notify_on_release to work in cgroups with OpenVZ?
