

Osmocom.org Servers - Bug #3071

docker-in-lxc setup results in high lxcfs CPU load

03/17/2018 12:23 PM - laforge

Status:	New	Start date:	03/17/2018
Priority:	Normal	Due date:	
Assignee:	laforge	% Done:	0%
Category:			
Target version:			
Spec Reference:			

Description

When looking at e.g. build2, I see the following entry in top:

```
686 root      20    0 2749704   5840   1328 S 370.5   0.0  15262:35 lxcfs
```

This means 370% CPU used on the lxcfs daemon.

When I strace lxcfs, it's always about the deb8build container, executing build jobs:

```
[pid 4254] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 17
[pid 4254] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 17
[pid 20990] openat(13, "./lxc/deb8build/memory.limit_in_bytes", O_RDONLY) = 16
[pid 20990] openat(13, "./lxc/memory.limit_in_bytes", O_RDONLY) = 16
[pid 20990] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 16
[pid 20990] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 16
[pid 20990] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 16
[pid 4140] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 18
[pid 551] openat(13, "./lxc/deb8build/memory.limit_in_bytes", O_RDONLY) = 20
[pid 551] openat(13, "./lxc/memory.limit_in_bytes", O_RDONLY) = 20
[pid 551] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 20
[pid 551] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 20
[pid 551] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 20
[pid 20991] openat(13, "./lxc/deb8build/memory.limit_in_bytes", O_RDONLY) = 22
[pid 20991] openat(13, "./lxc/memory.limit_in_bytes", O_RDONLY) = 22
[pid 20991] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 22
[pid 20991] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 22
[pid 20991] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 22
[pid 3976] openat(13, "./lxc/deb8build/memory.memsw.limit_in_bytes", O_RDONLY) = -1 ENOENT (No such file or directory)
[pid 15513] openat(13, "./lxc/deb8build/memory.limit_in_bytes", O_RDONLY) = 19
[pid 15513] openat(13, "./lxc/memory.limit_in_bytes", O_RDONLY) = 19
[pid 15513] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 19
[pid 15513] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 19
[pid 15513] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 19
[pid 4140] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 18
[pid 23069] openat(13, "./lxc/deb8build/memory.limit_in_bytes", O_RDONLY) = 21
[pid 23069] openat(13, "./lxc/memory.limit_in_bytes", O_RDONLY) = 21
[pid 23069] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 21
[pid 23069] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 21
[pid 23069] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 21
[pid 4302] openat(13, "./lxc/deb8build/memory.limit_in_bytes", O_RDONLY) = 23
[pid 4302] openat(13, "./lxc/memory.limit_in_bytes", O_RDONLY) = 23
[pid 4302] openat(13, "./memory.limit_in_bytes", O_RDONLY) = 23
[pid 4302] openat(13, "./lxc/deb8build/memory.usage_in_bytes", O_RDONLY) = 23
[pid 4302] openat(13, "./lxc/deb8build/memory.stat", O_RDONLY) = 23
```

The current theory is that dockers libcontainer is **heavily** reading cgroups (<https://github.com/docker/libcontainer/blob/master/cgroups/fs/memory.go>), probably to get memory usage statistics.

As cgroups is not provided by the kernel on the host OS, but emulated via FUSE + lxcfs, those frequent read accesses are causing

an enormous load on the lxcfs daemon.

Now the question is:

- can we live without lxcfs ?
- can we disable docker from reading all those [unused] stats all the time
- do we have to use docker natively on the machine (deb9), rather than inside containers (deb8, deb9, ...)?

History

#1 - 03/17/2018 12:24 PM - laforge

- Description updated