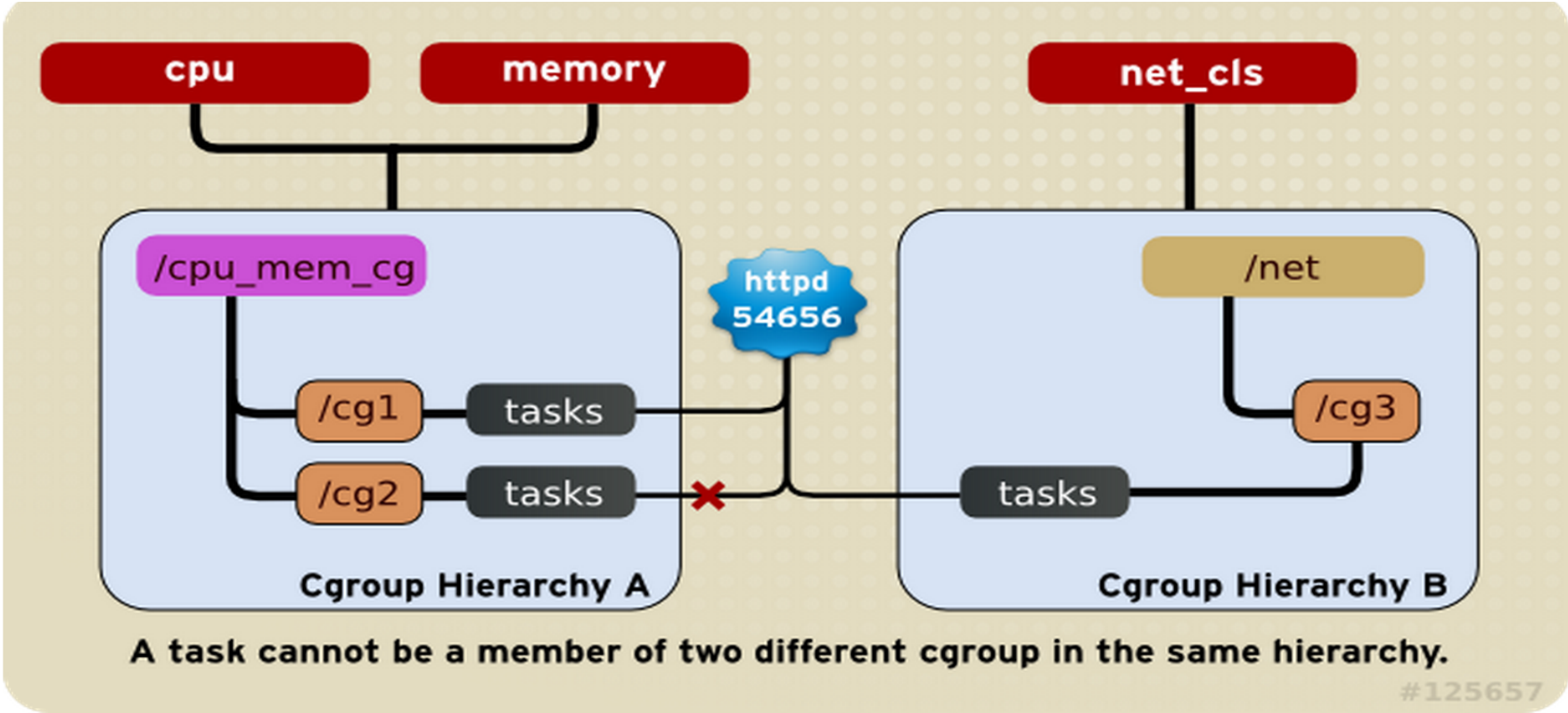
Cgroups and Namespaces

- Madhurranjan

Cgroups

- Allocate resources
- Manage resources
- 11 cgroup subsystems
- Popular ones : cpu , memory , blkio , cpuacct

Cgroups - hierarchy, cgroups , tasks (source : Redhat documentation)



Managing Cgroups

- libcgroup - Rhel 6
  - cgcreate
  - cgexec
  - /etc/cgroup.conf
- systemd (fedora)
- We’ll look at both of them :)

Cgroup hierarchy

- mkdir /cgroup/foo

mount -t cgroup -o memory,cpu foo /cgroup/foo/

(subsystems - memory and cpu)

- Can’t mount memory separately now.

mount -t cgroup -o memory foo3 /cgroup/foo3/

mount: foo3 already mounted or /cgroup/foo3/ busy

Lets create CGgroups now!

cgcreate -g cpu:A

cgcreate -g cpu:B

cgcreate -g cpu:C

Mandatory -

mkdir /cgroup/cpu_and_mem

mount -t cgroup -o memory,cpu cpu_and_mem /cgroup/cpu_and_mem/

echo 0 >> /cgroup/cpu_and_mem/C/cpuset.cpus

echo 0 >> /cgroup/cpu_and_mem/C/cpuset.mems

Cpu share example

```
sudo cgexec -g cpu:C dd if=/dev/zero
of=/dev/null &

sudo cgexec -g cpu:A dd if=/dev/zero
of=/dev/null &
```

Out of memory example:

echo 1M > /cgroup/cpu_and_mem/A/memory.limit_in_bytes

cat /cgroup/cpu_and_mem/A/memory.limit_in_bytes

service httpd start

echo <pid> >> /cgroup/cpu_and_mem/A/tasks

ab -n 1000000 -c 1000 http://127.0.0.1/sample.html

Cgsnapshot

cgsnapshot -s > cgconfig-example.conf

Docker - where do I see cgroups

ls /sys/fs/cgroup/< subsystem >

Simple perf test with docker?

echo 1M > /sys/fs/cgroup/memory/system.slice/docker-<>.scope/memory.limit_in_bytes

ab -n 10000 -c 1000 http://<ip>/sample.html

Namespaces

Abstraction to make sure system resources seem to the processes that run within namespaces like they are isolated

6 Namespaces

- Mount namespace : 2.4.19
- UTS namespace: 2.6.19
- IPC namespace: 2.6.19
- PID namespace: 2.6.24
- Network namespace: 2.6.29
- User namespace: 3.8 *

* unprivileged containers

Where do I find it ?

ls /proc/<pid>/ns/

Shows what namespaces were used or implemented

UTS namespace

- unshare --help
- unshare -u /bin/bash
- hostname
- uname -a

PID Namespace : Docker process

Process outside and inside have different pids and are mapped

Network namespace

- Add 2 network namespaces:

  ip netns add myns1

  ip netns add myns2

Network namespace

ip link set eth1 netns myns2

ip netns exec myns2 bash

ip netns exec myns2 bash

ip link set eth1 netns 1

Time remaining:

ip link add name **if_one** type veth peer name **if_one**_peer

ip link set dev **if_one**_peer netns myns2

ifconfig if_one 10.0.0.1 up

ifconfig if_one_peer 10.0.0.2 up

Docker : Network namespace - contd

ln -s /proc/<pid>/ns/net /var/run/netns/<name>

ip netns exec <symlink> netstat -i

ab -n 10 -c 1 http://172.17.0.2/sample.html

ip netns exec <name> bash

Looking at the inodes associated

ip netns exec myns1 ls -la /proc/self/ns/

total 0

dr-x--x--x. 2 root root 0 Sep 20 06:38 .

dr-xr-xr-x. 8 root root 0 Sep 20 06:38 ..

lrwxrwxrwx. 1 root root 0 Sep 20 06:38 ipc -> ipc:[4026531839]

lrwxrwxrwx. 1 root root 0 Sep 20 06:38 mnt -> mnt:[4026532257]

lrwxrwxrwx. 1 root root 0 Sep 20 06:38 net -> net:[4026532198]

lrwxrwxrwx. 1 root root 0 Sep 20 06:38 pid -> pid:[4026531836]

lrwxrwxrwx. 1 root root 0 Sep 20 06:38 uts -> uts:[4026531838]

Useful Links

Dr Rami Rosen’s presentations : http://ramirose.wix.com/ramirosen

lwn.net articles on namespaces(lwn.net/Articles/531114/)

cgroups - redhat documentation:

https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Resource_Management_Guide/ch01.html

Thank you !