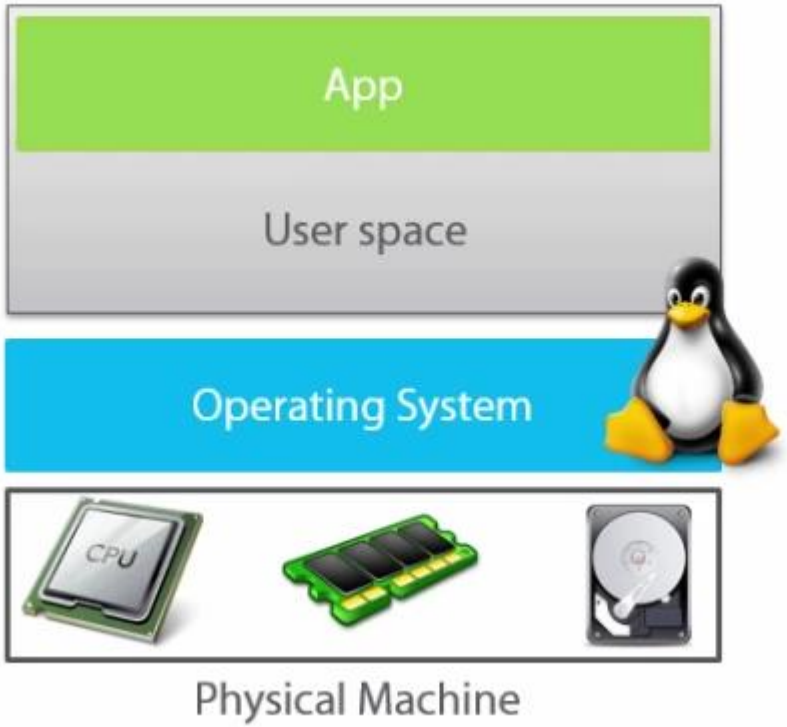
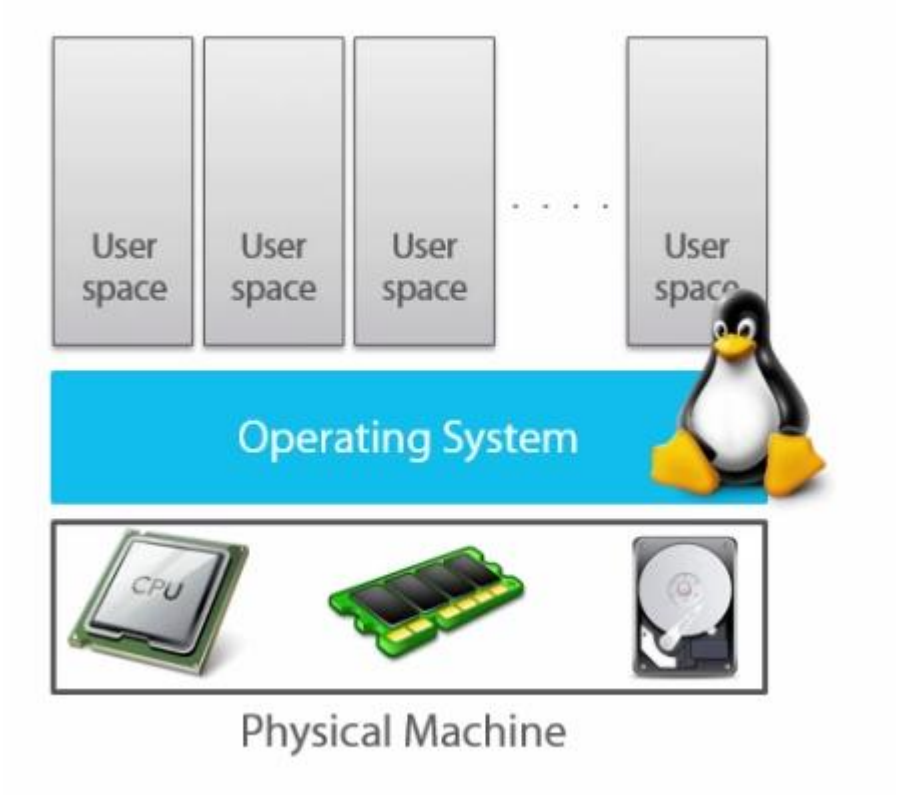
How Containers Work

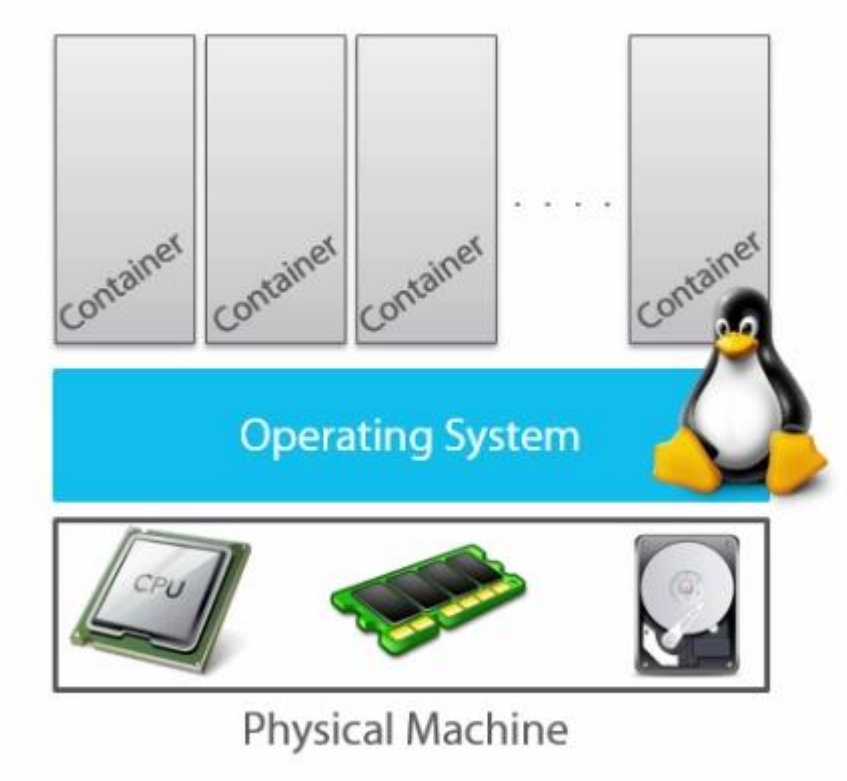
Physical Machine



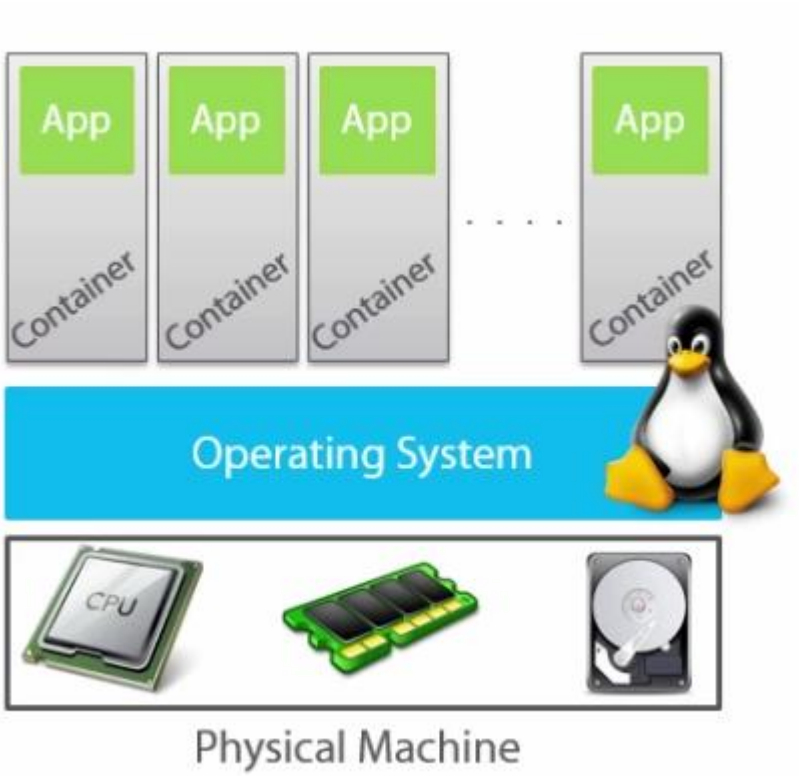
What containers Do?



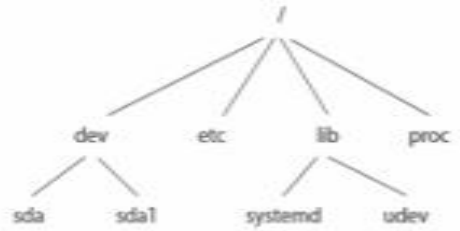
Each isolated User space is container

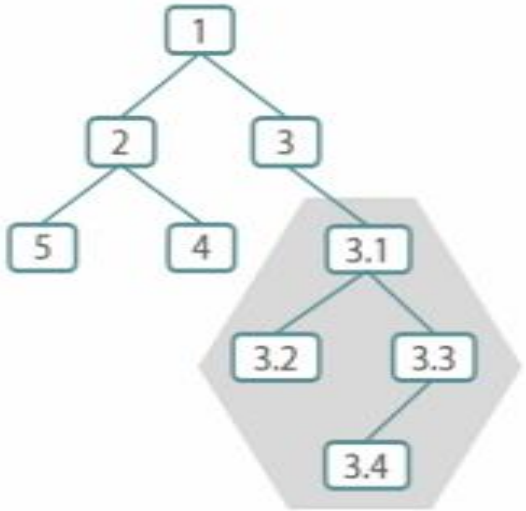


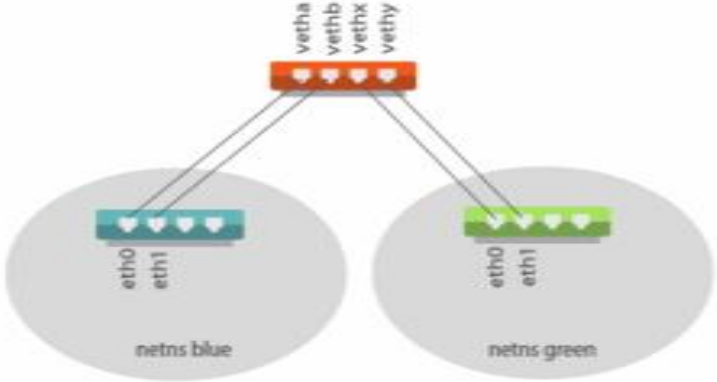
In each container we install App

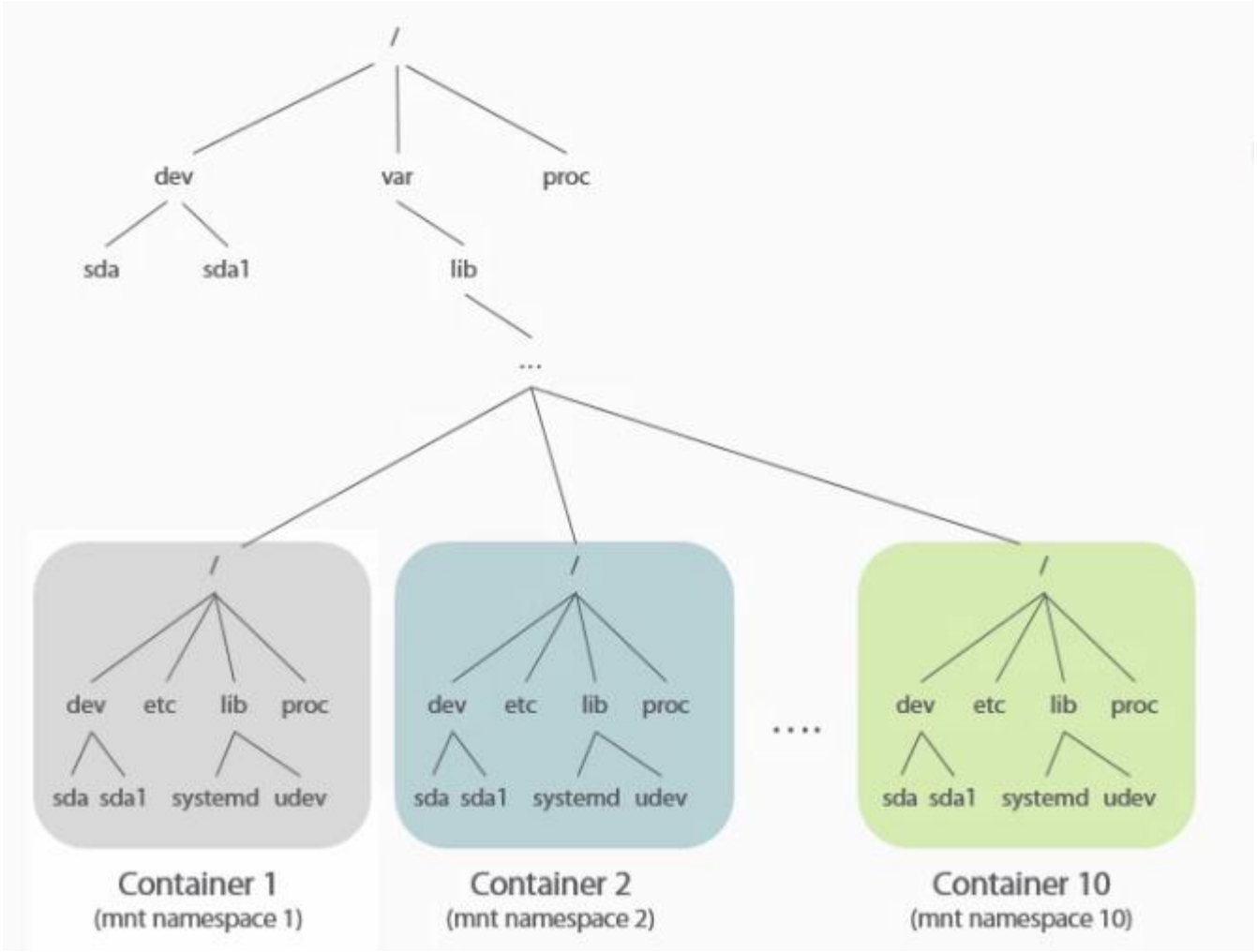


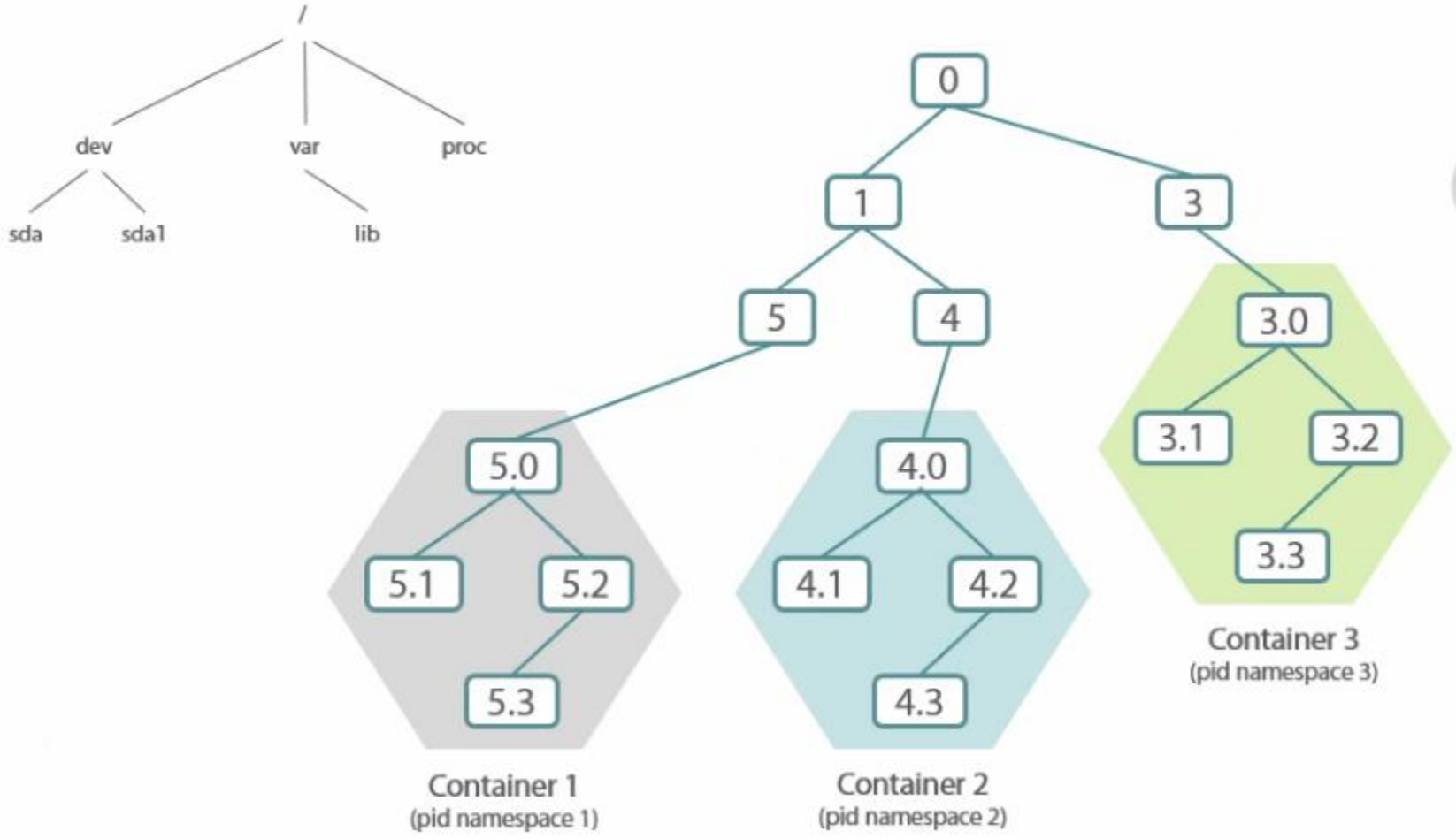
Process Tree, File System, Network

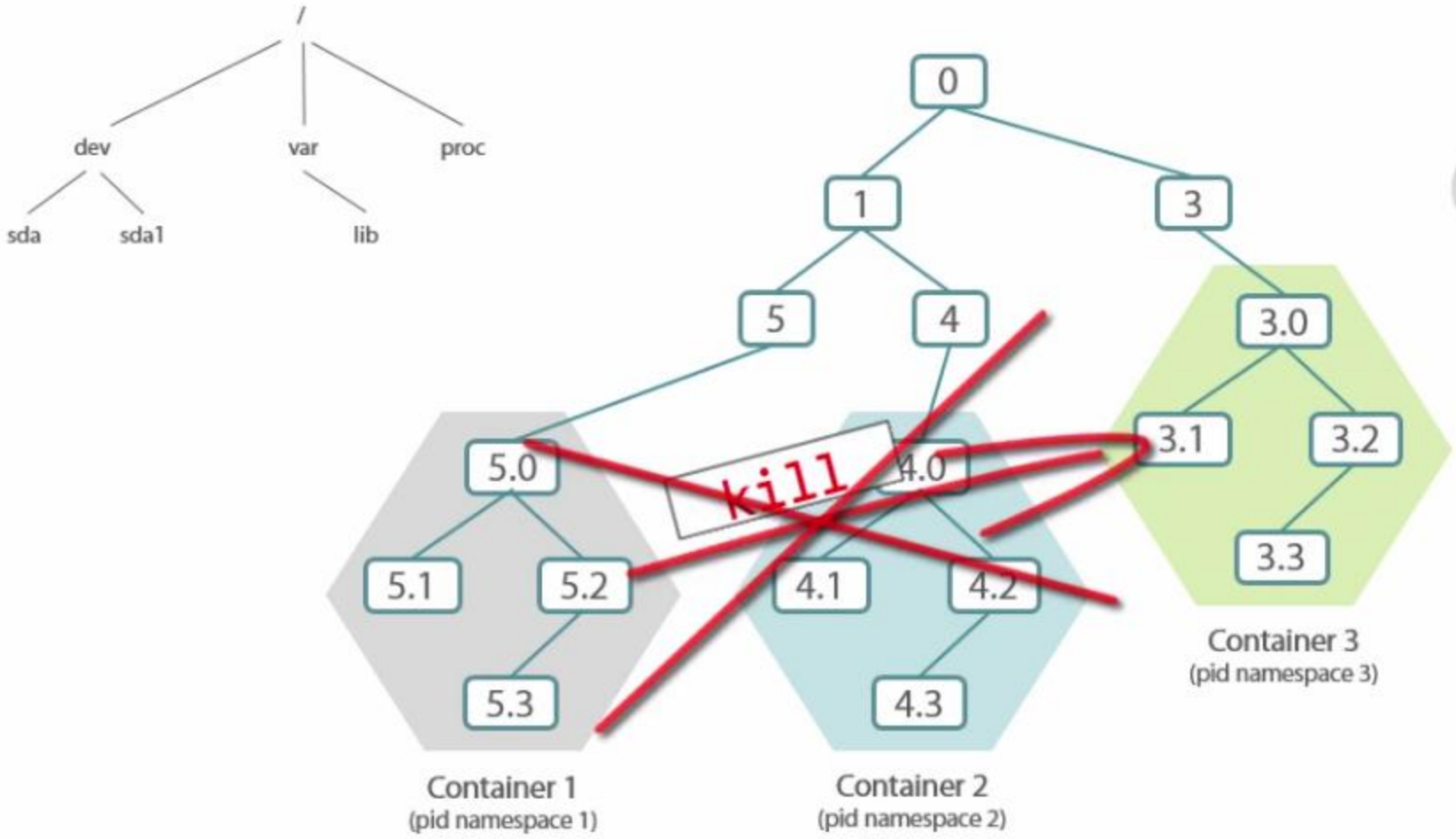







/
dev
var
proc
sda
sda1
lib
...
/
dev
etc
lib
proc
sda
sda1
systemd
udev
Container 1
(mnt namespace 1)
/
dev
etc
lib
proc
sda
sda1
systemd
udev
Container 2
(mnt namespace 2)
....
/
dev
etc
lib
proc
sda
sda1
systemd
udev
Container 10
(mnt namespace 10)

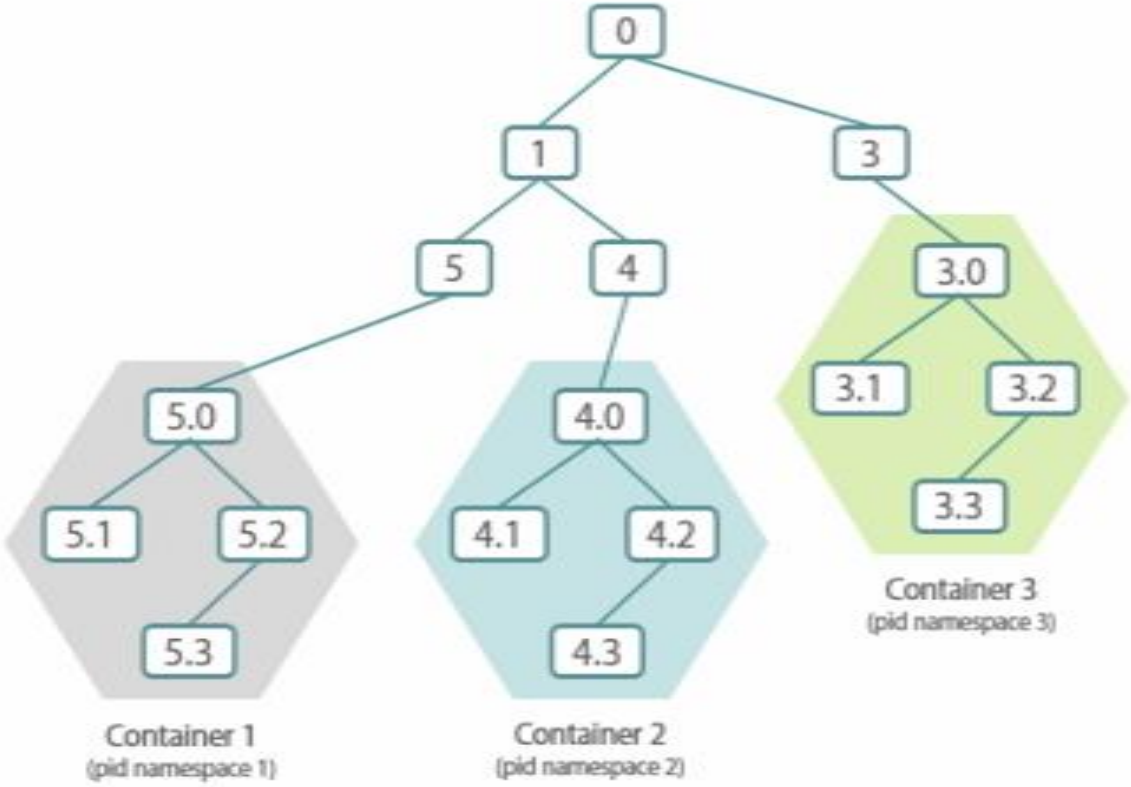

/
dev
var
proc
sda
sda1
lib
0
1
3
5
4
3.0
3.1
3.2
3.3
5.0
5.1
5.2
5.3
4.0
4.1
4.2
4.3
Container 1
(pid namespace 1)
Container 2
(pid namespace 2)
Container 3
(pid namespace 3)


/
dev
var
proc
sda
sda1
lib
0
1
3
5
4
3.0
5.0
4.0
3.1
3.2
kill
5.1
5.2
4.1
4.2
3.3
5.3
4.3
Container 1
(pid namespace 1)
Container 2
(pid namespace 2)
Container 3
(pid namespace 3)

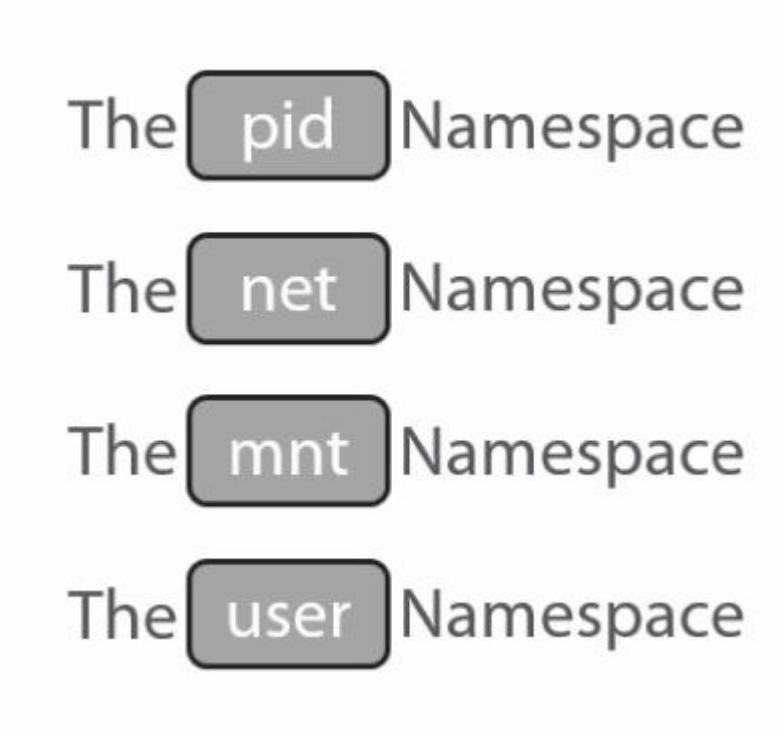
How are these isolations created ???

Kernel Namespaces

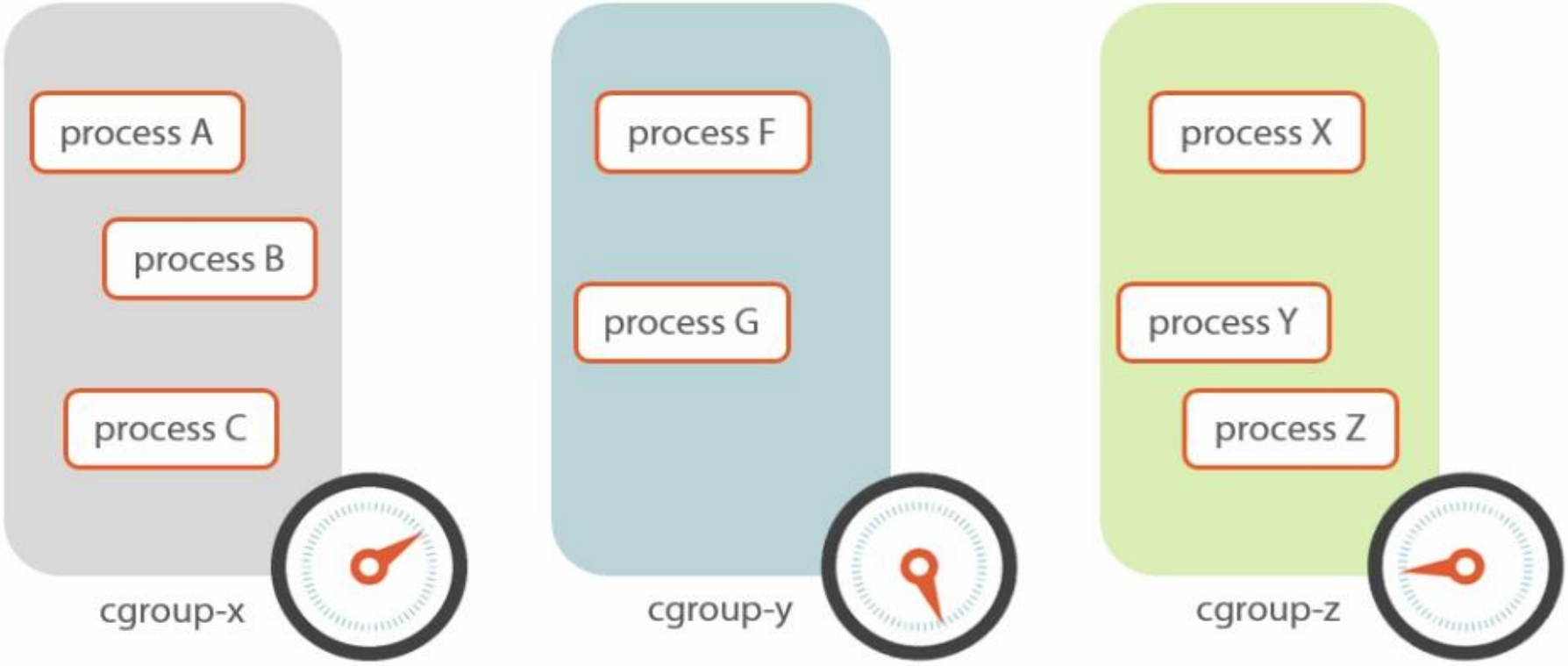
The pid Namespace



Only Process Namespace



Control Groups (cgroups)



Capabilities

root

CAP_AUDIT_CONTROL ✓

CAP_CHOWN ✓

CAP_DAC_OVERRIDE ✓

CAP_KILL ✓

CAP_NET_BIND_SERVICE ✓

CAP_SETUID ✓

. . . . . ✓

Capabilities

root

CAP_AUDIT_CONTROL ✓

CAP_CHOWN ✓

CAP_DAC_OVERRIDE ✓

CAP_KILL ✓

CAP_NET_BIND_SERVICE ✓

CAP_SETUID ✓

. . . . . ✓