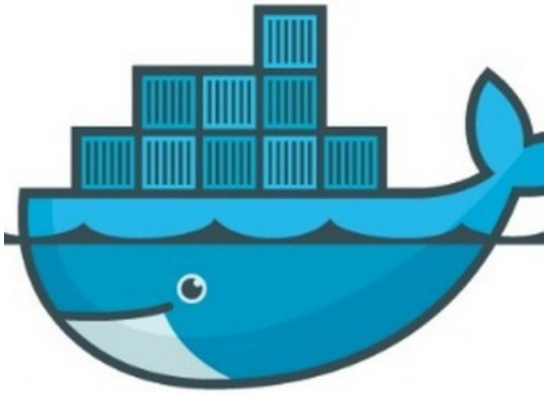


“Containers do not contain”

(orig. by Dan Walsh)

Maciej Lasyk

Devopsdays Warsaw

2015-11-25

Kelsey Hightower @kelseyhightower · 21h

SELinux is being proposed for securing Linux containers. Has usability improved that much to stop people from turning it off by default?

9 26

Kelsey Hightower @kelseyhightower · 21h

SELinux is being proposed for securing Linux containers. Has usability improved that much to stop people from turning it off by default?

9 26

Maciek Lasyk @docent_net · 20h

@kelseyhightower actually I was able to explain #selinux basics2my 6 years old kid w/ coloring book (cc: @rhatdan) people.redhat.com/duffy/selinux/...

How does security look like?

And seriously...

Do you know this guy?

And seriously...

Do you know this guy?

So he has something to tell you...

http://www.youtube.com/watch?v=o5snlP8Y5GY

Linux OS security

DAC (Discretionary access control)

basic ACLs

chmod

Linux OS security – extended ACLs

DAC (Discretionary access control)

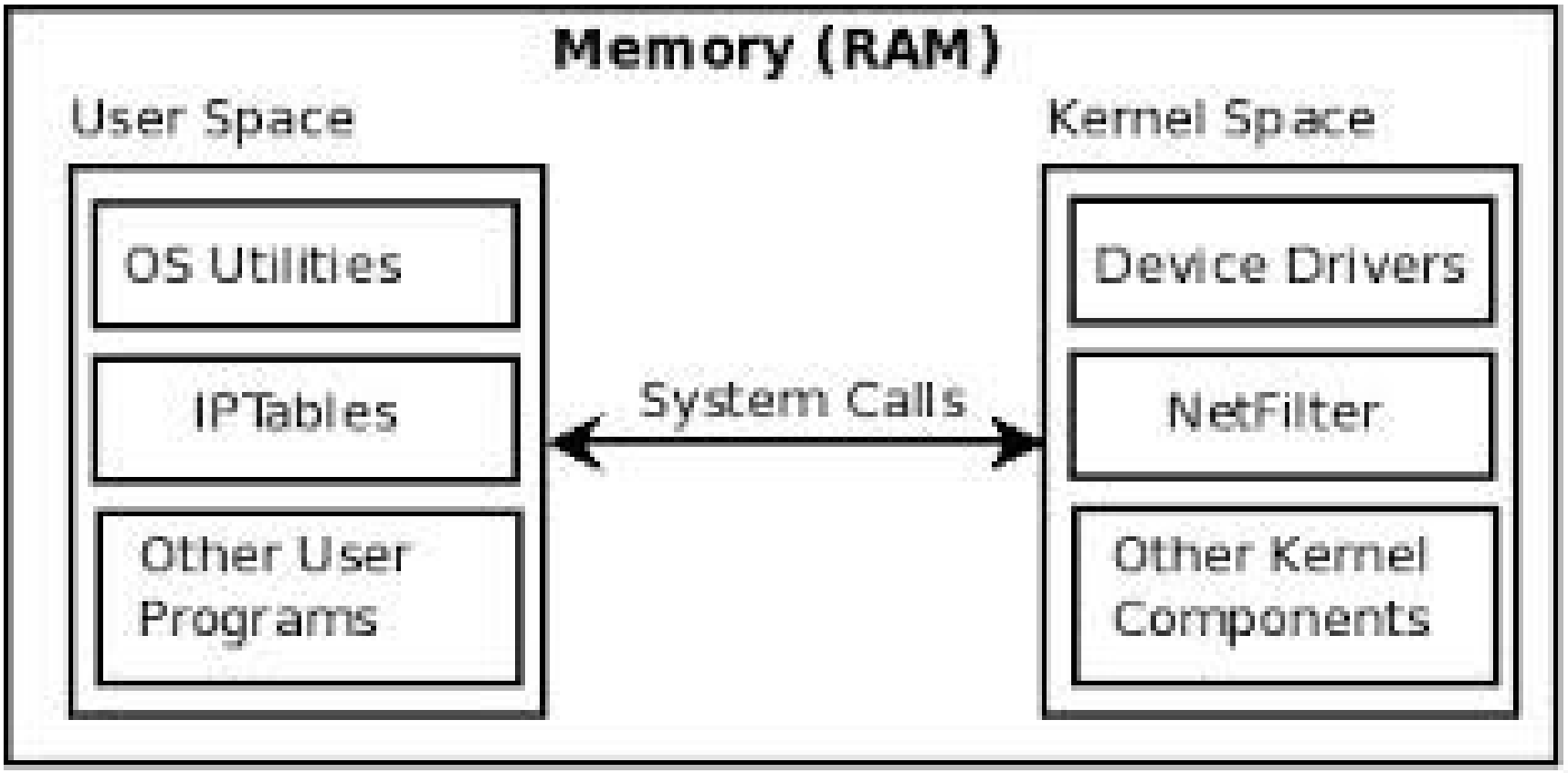
setfacl / getfacl

Linux OS security – Linux Security Modules

MAC (Mandatory Access Control)

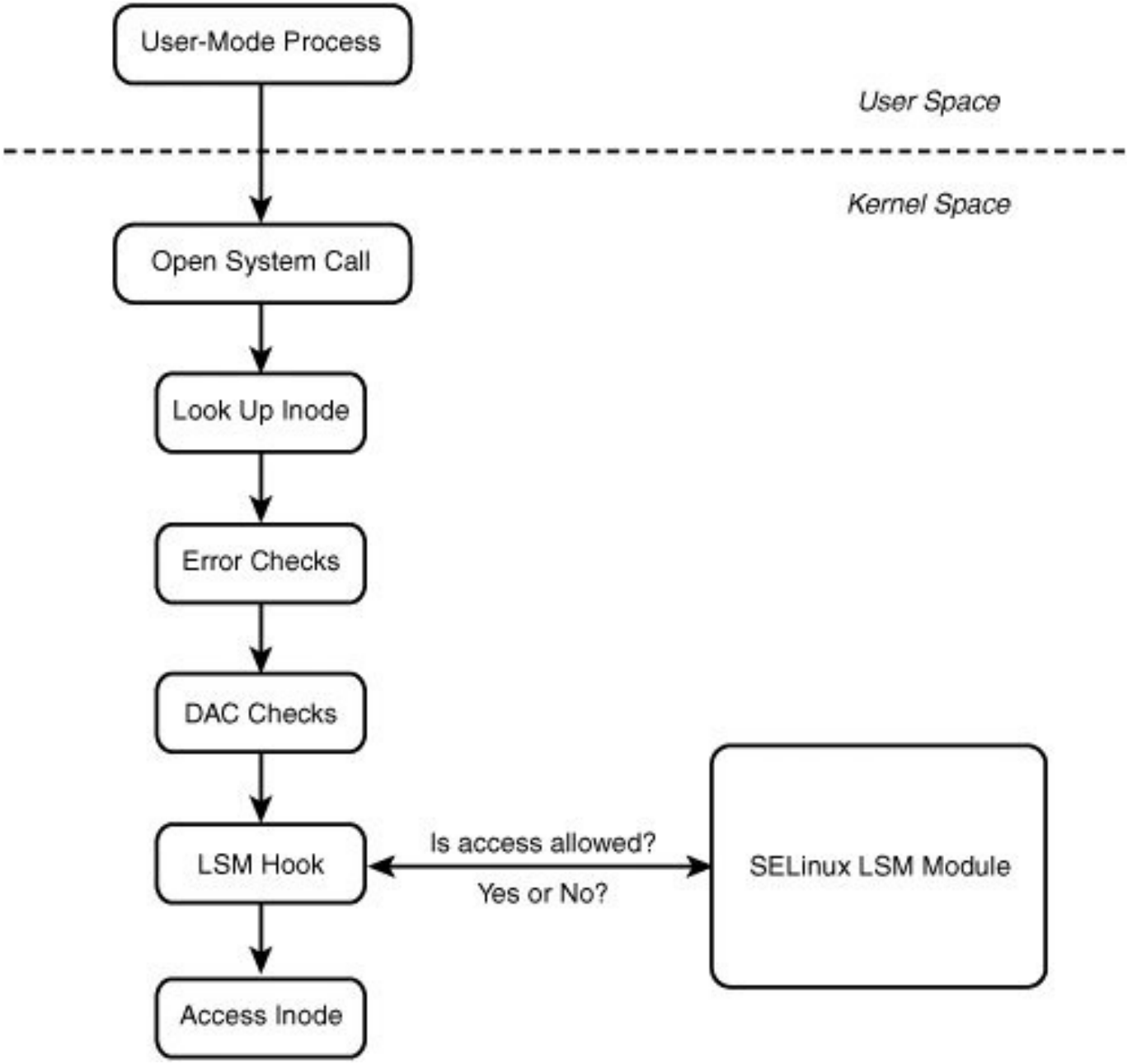
(LSMs)

SELinux – how it works?

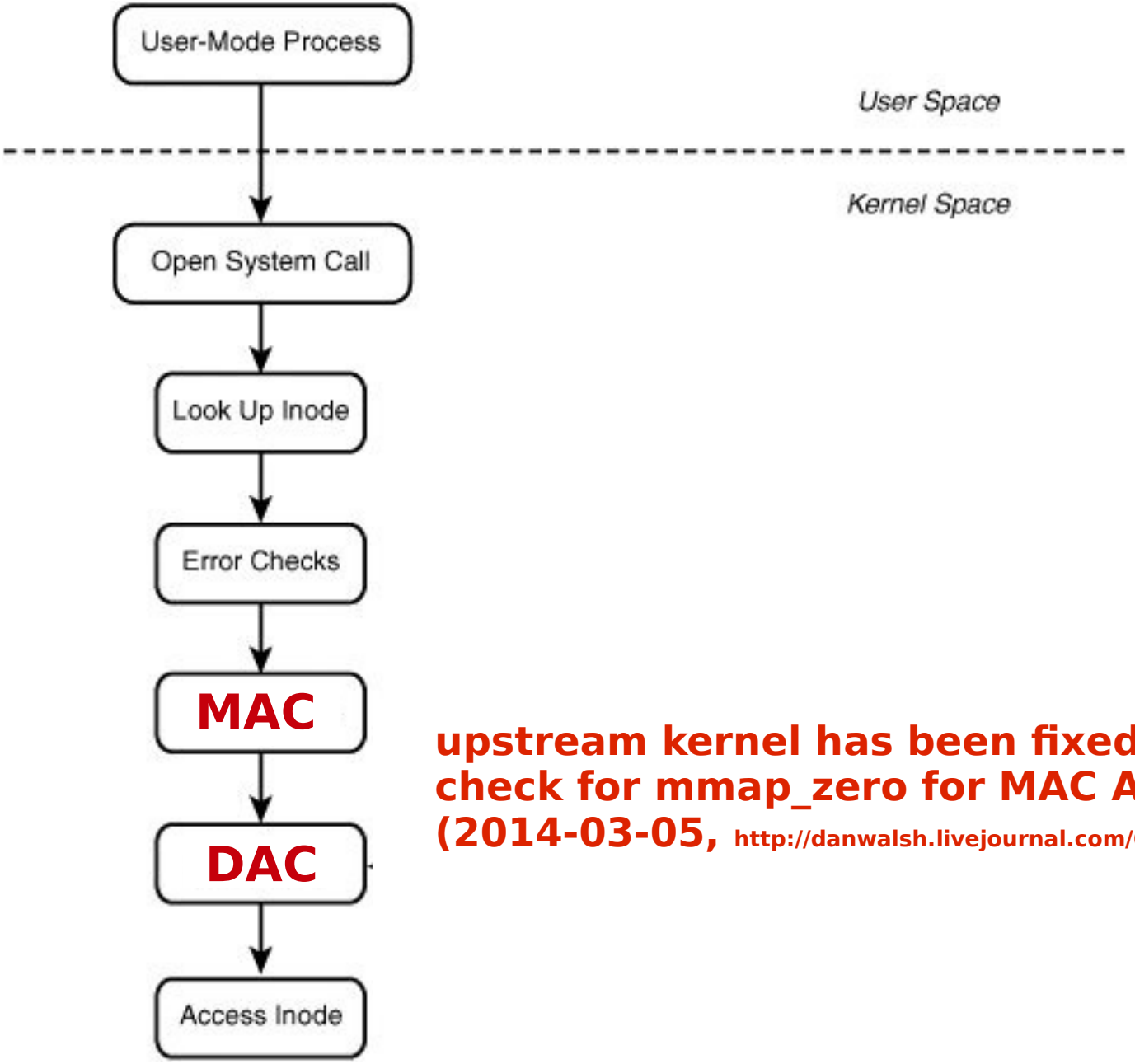


syscalls work like interfaces for accessing some resources

SELinux – how it works?

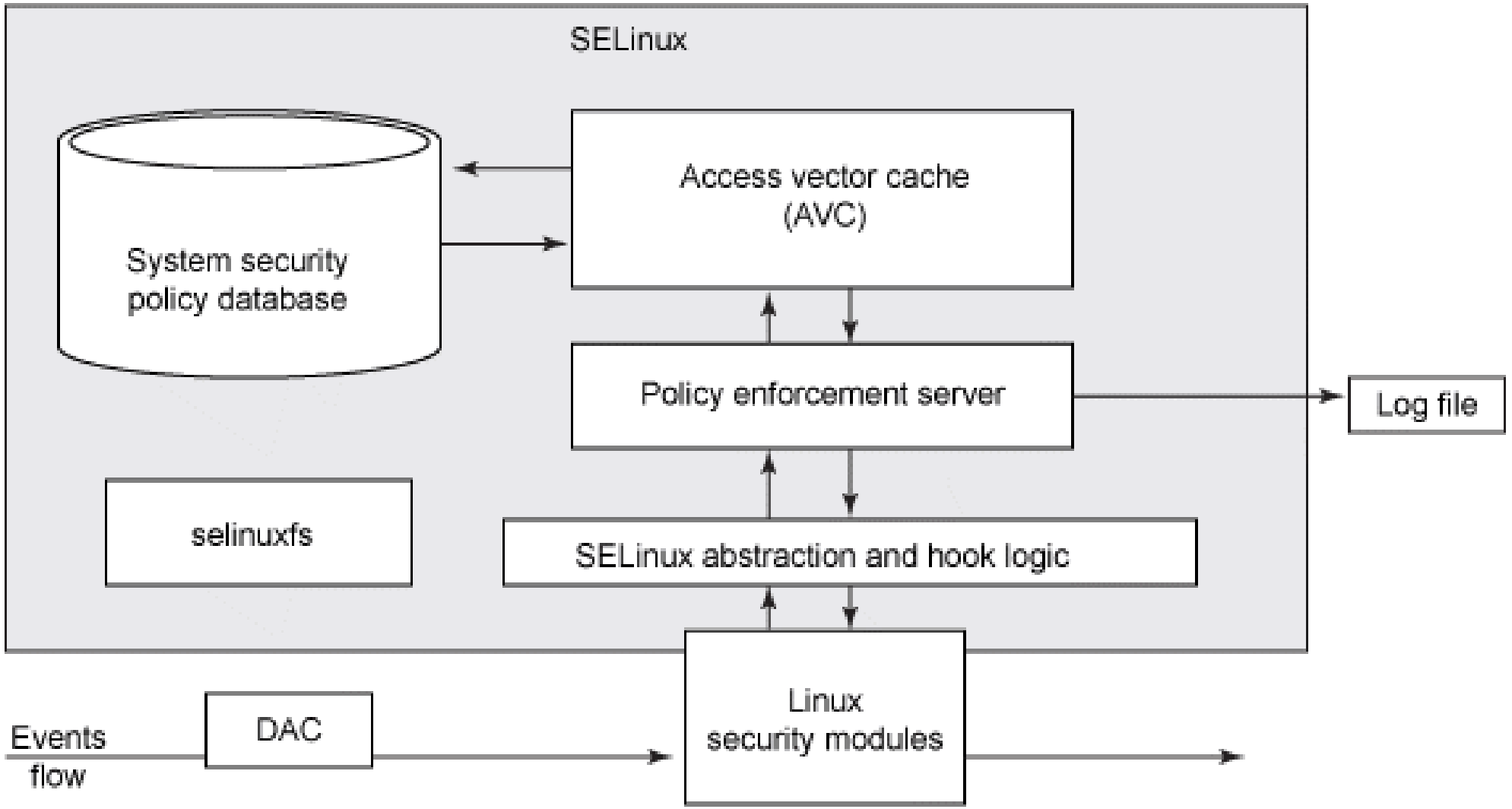


SELinux – how it works?



upstream kernel has been fixed to report check for mmap_zero for MAC AFTER DAC (2014-03-05, http://danwalsh.livejournal.com/69035.html)

SELinux – how it works?



SELinux – performance

- http://www.nsa.gov/research/_files/selinux/papers/freenix01/node18.shtml#sec:perf:macro

Table: Macrobenchmark results. The elapsed and system times for a ``time make'' on the Linux 2.4.2 kernel sources are shown in minutes and seconds. The latency in seconds and throughput in MBits per second are shown for the WebStone benchmark.

	Base	SELinux	Overhead
elapsed	11:14	11:15	0%
system	00:49	00:51	4%
latency	0.56	0.56	0%
throughput	8.29	8.28	0%

Just test it yourself: git://git.selinuxproject.org/~serge/selinux-testsuite

SELinux – performance

- http://www.nsa.gov/research/_files/selinux/papers/freenix01/node18.shtml#sec:perf:macro

Table: Macrobenchmark results. The elapsed and system times for a ``time make'' on the Linux 2.4.2 kernel sources are shown in minutes and seconds. The latency in seconds and throughput in MBits per second are shown for the WebStone benchmark.

	Base	SELinux	Overhead
elapsed	11:14	11:15	0%
system	00:49	00:51	4%
latency	0.56	0.56	0%
throughput	8.29	8.28	0%

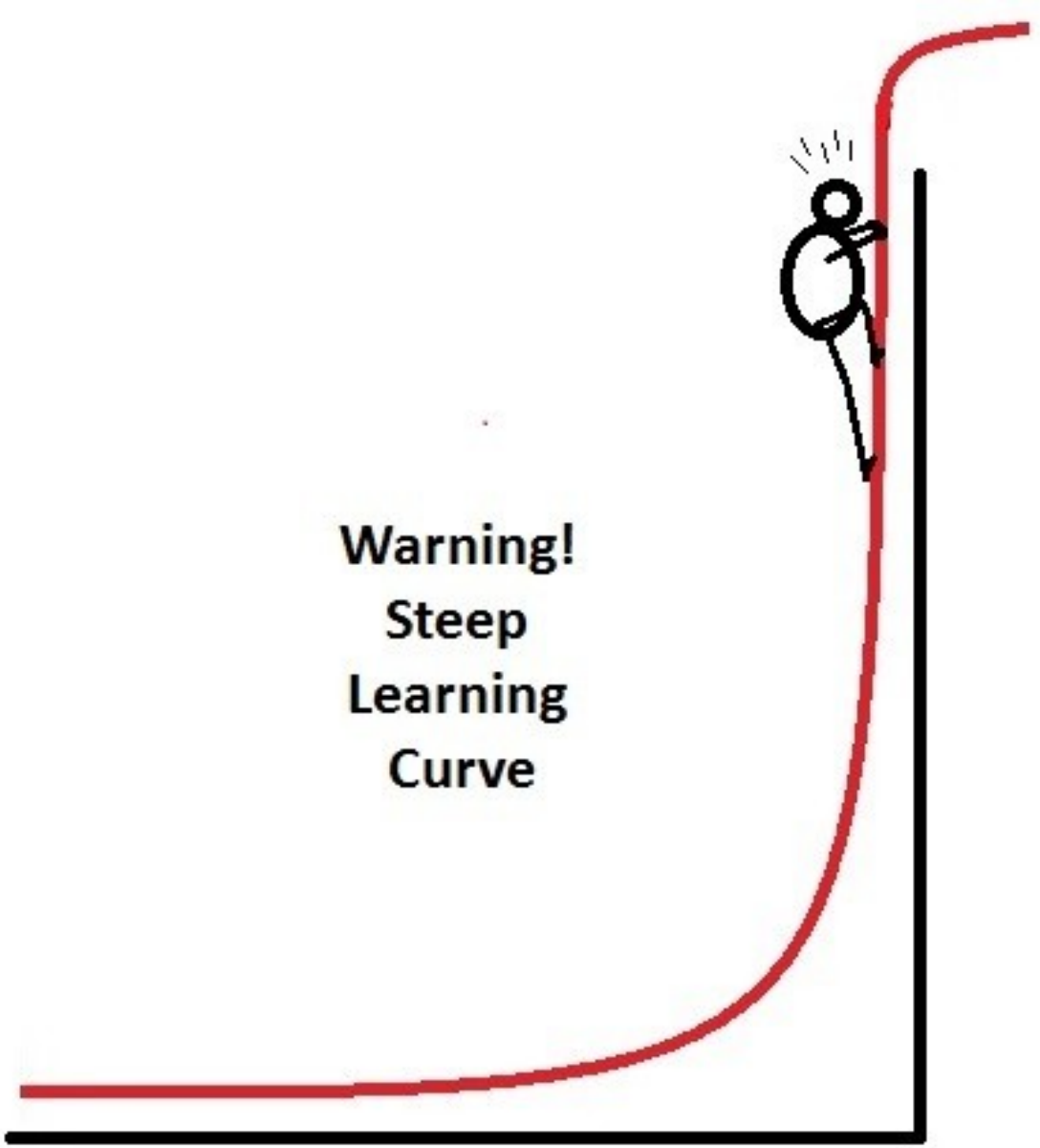
avcstat

uptime: 10h

hit ratio: 99.94%!
(57mln of lookups)

Just test it yourself: git://git.selinuxproject.org/~serge/selinux-testsuite

SELinux – learning curve



SELinux and...

SELinux and Android

- from 4.3 – permissive
- from 4.4 enforcing
- Will help us with BYOD :)
- No setuid/setgid programs (4.3)

http://selinuxproject.org/page/SEAndroid

http://source.android.com/devices/tech/security/se-linux.html

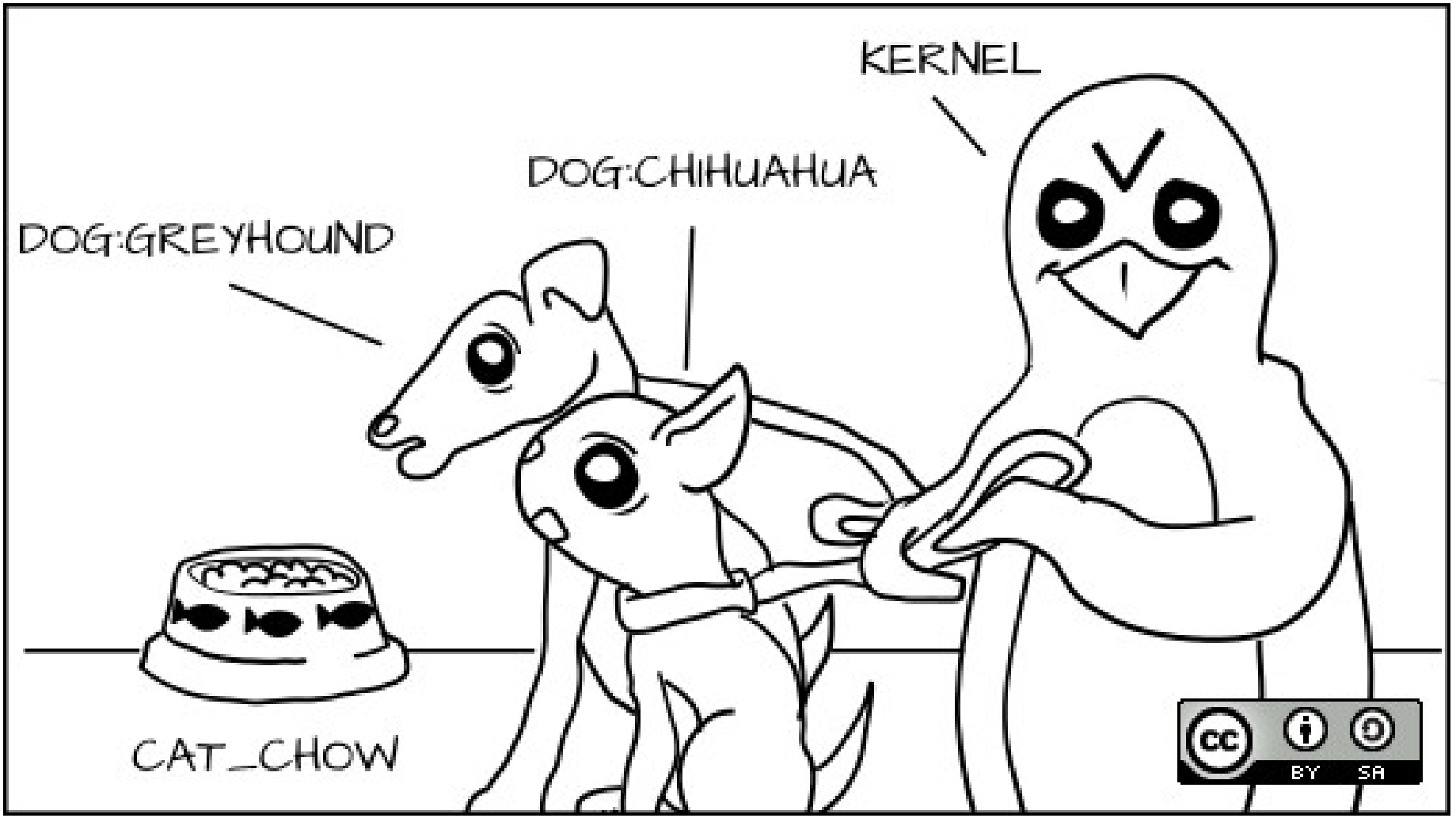
SELinux primer

stopdisablingselinux.com

or

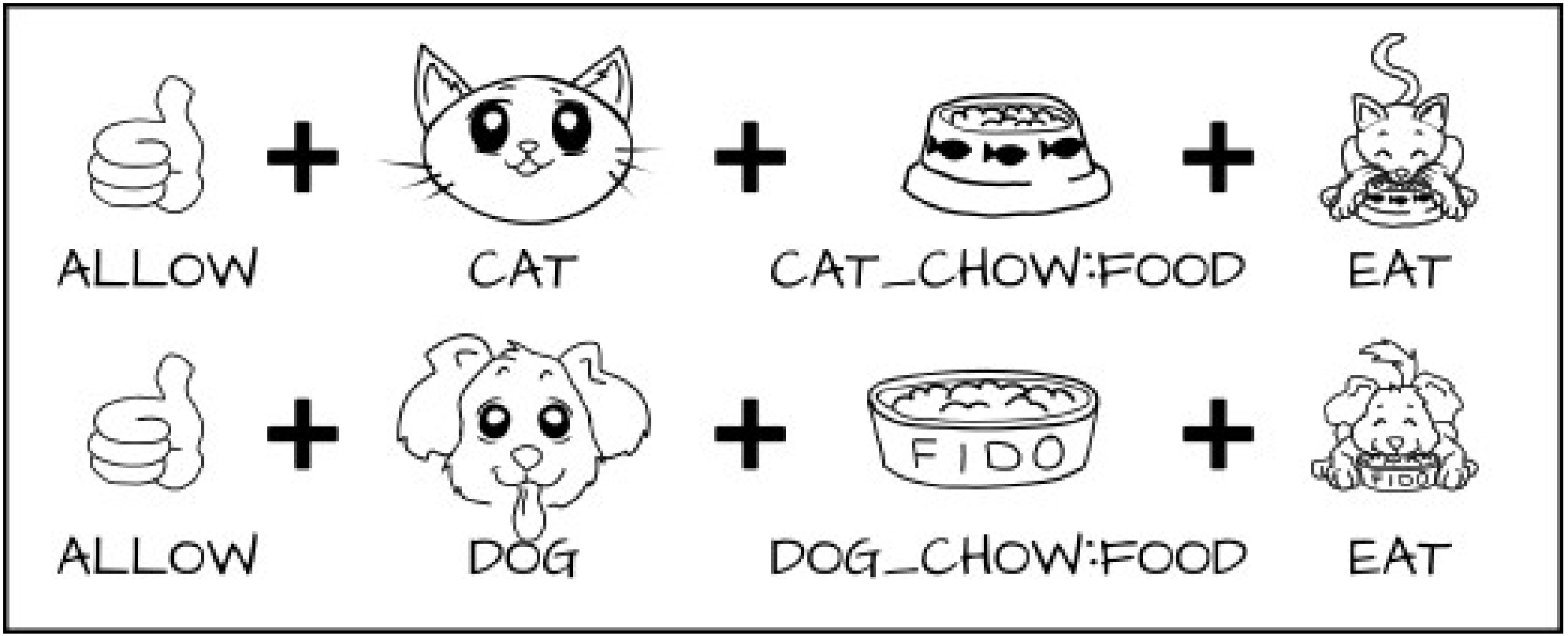
http://opensource.com/business/13/11/selinux-policy-guide

SELinux primer



Everyone gets a label!

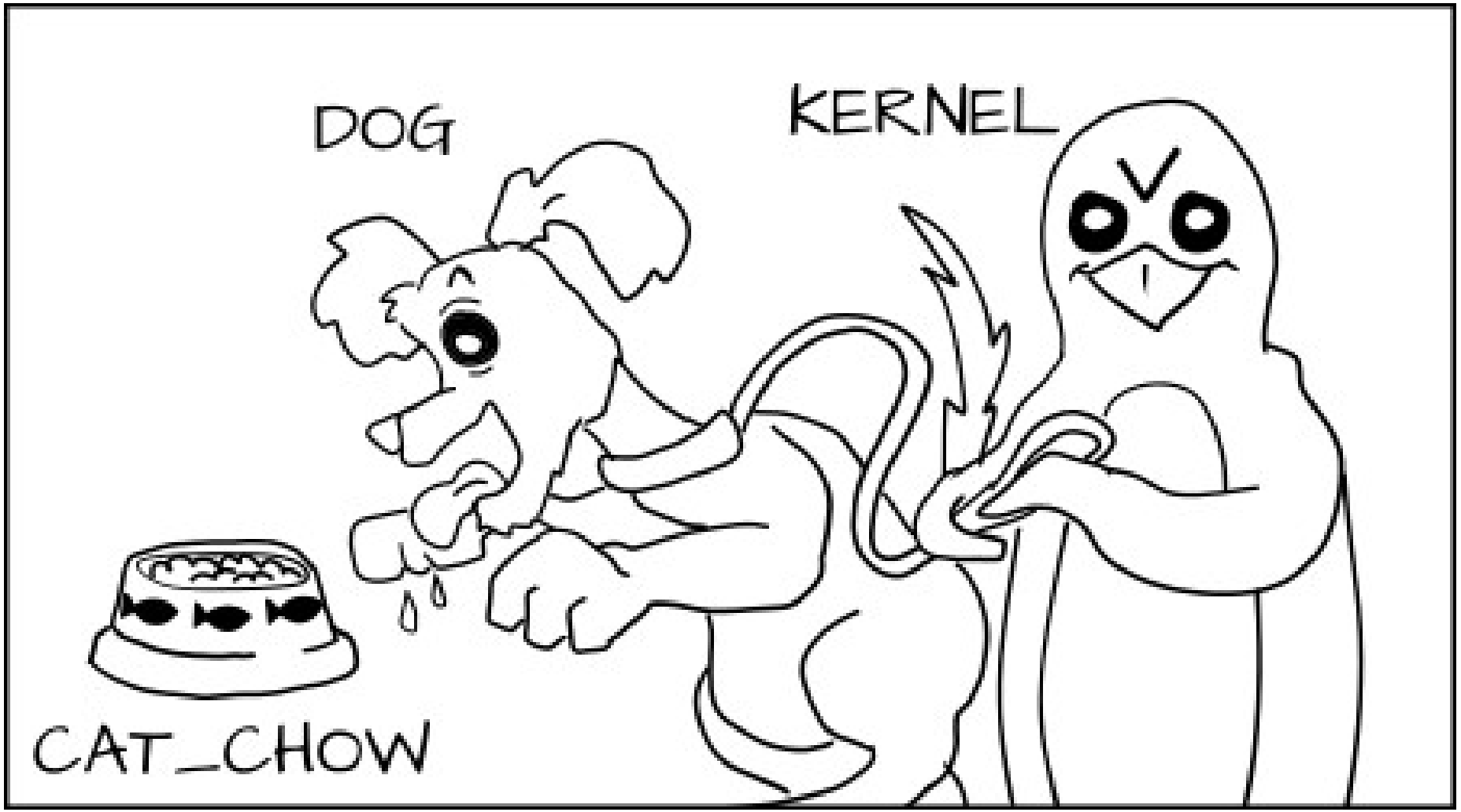
SELinux primer



allow cat cat_chow:food eat;

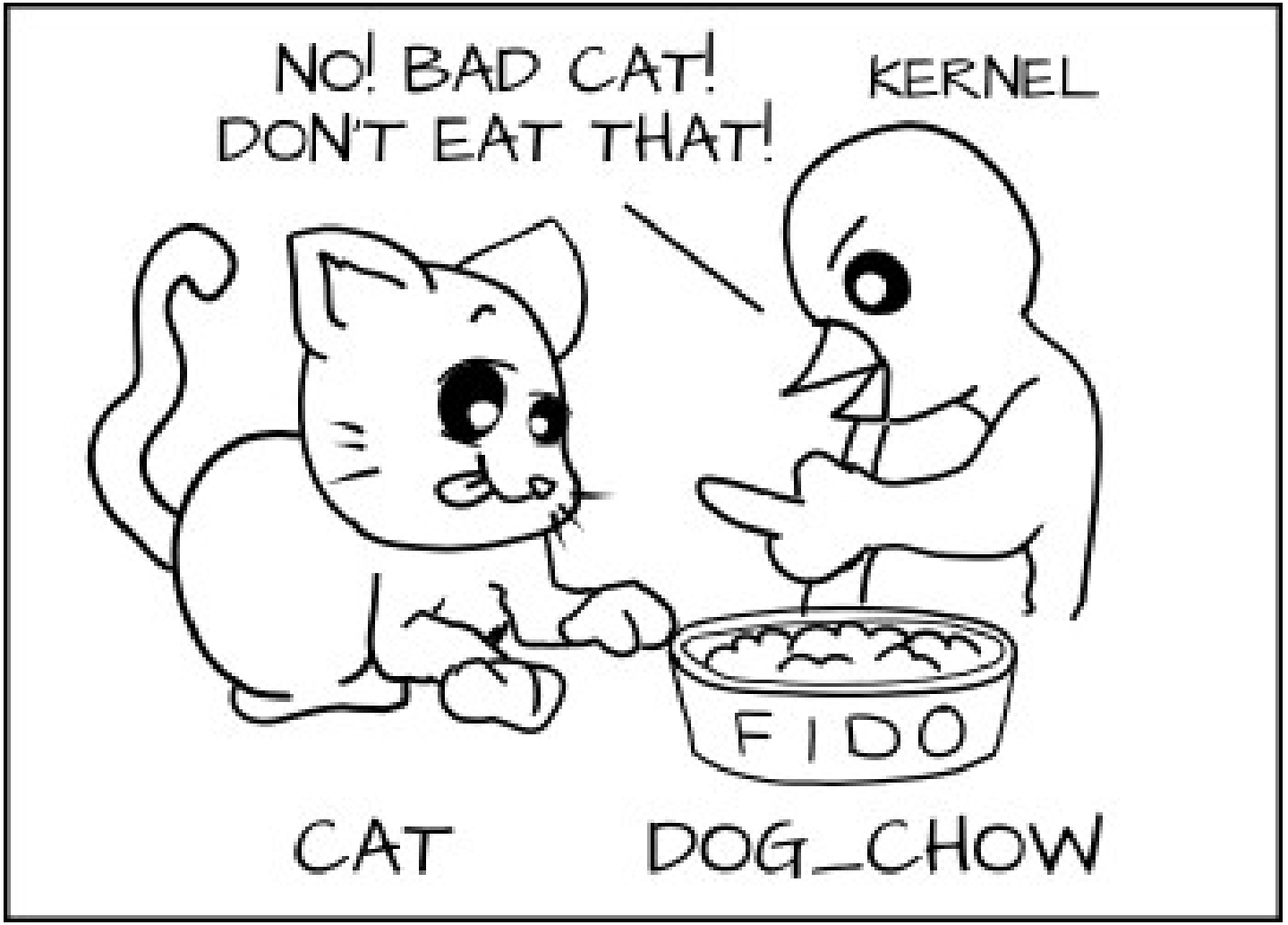
allow dog dog_chow:food eat;

SELinux primer



AVC (Access Vector Cache)

SELinux primer



AVC (Access Vector Cache)

SELinux primer

In real world...

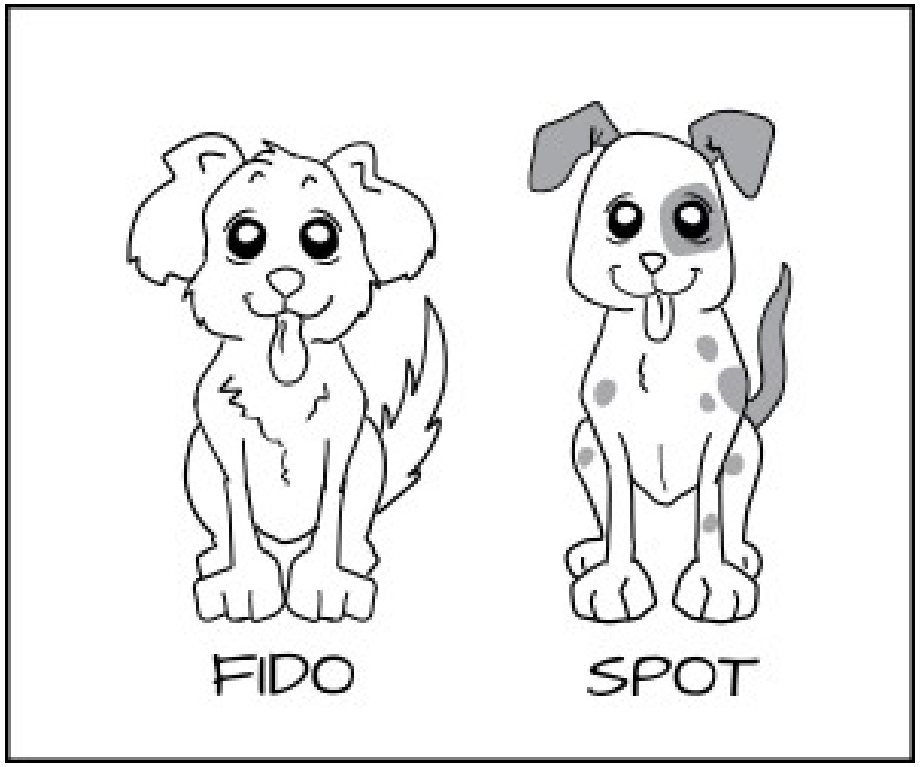
process: httpd_t

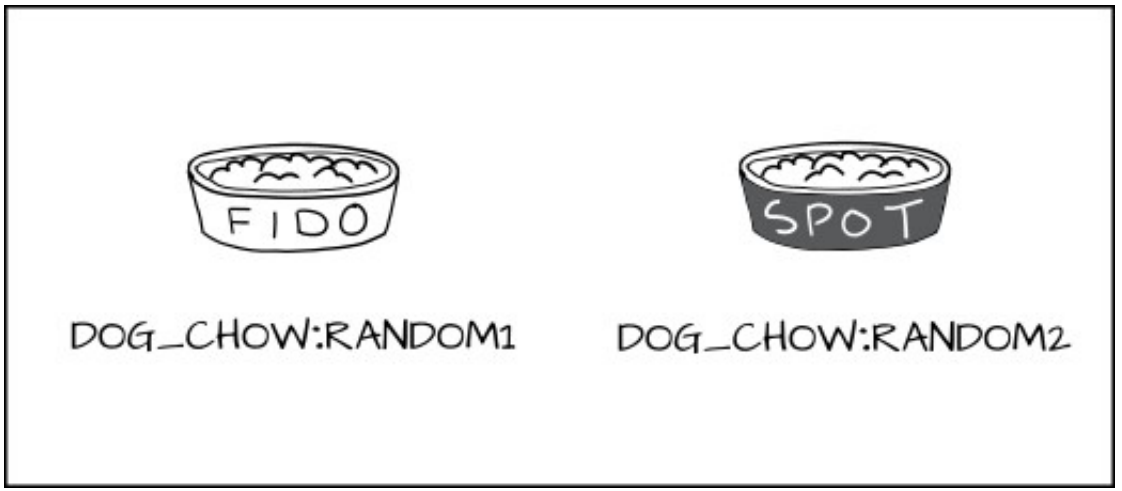
files under Apache: httpd_sys_content_t

database data: mysqld_data_t

hacked Apache process can not access mysqld files!

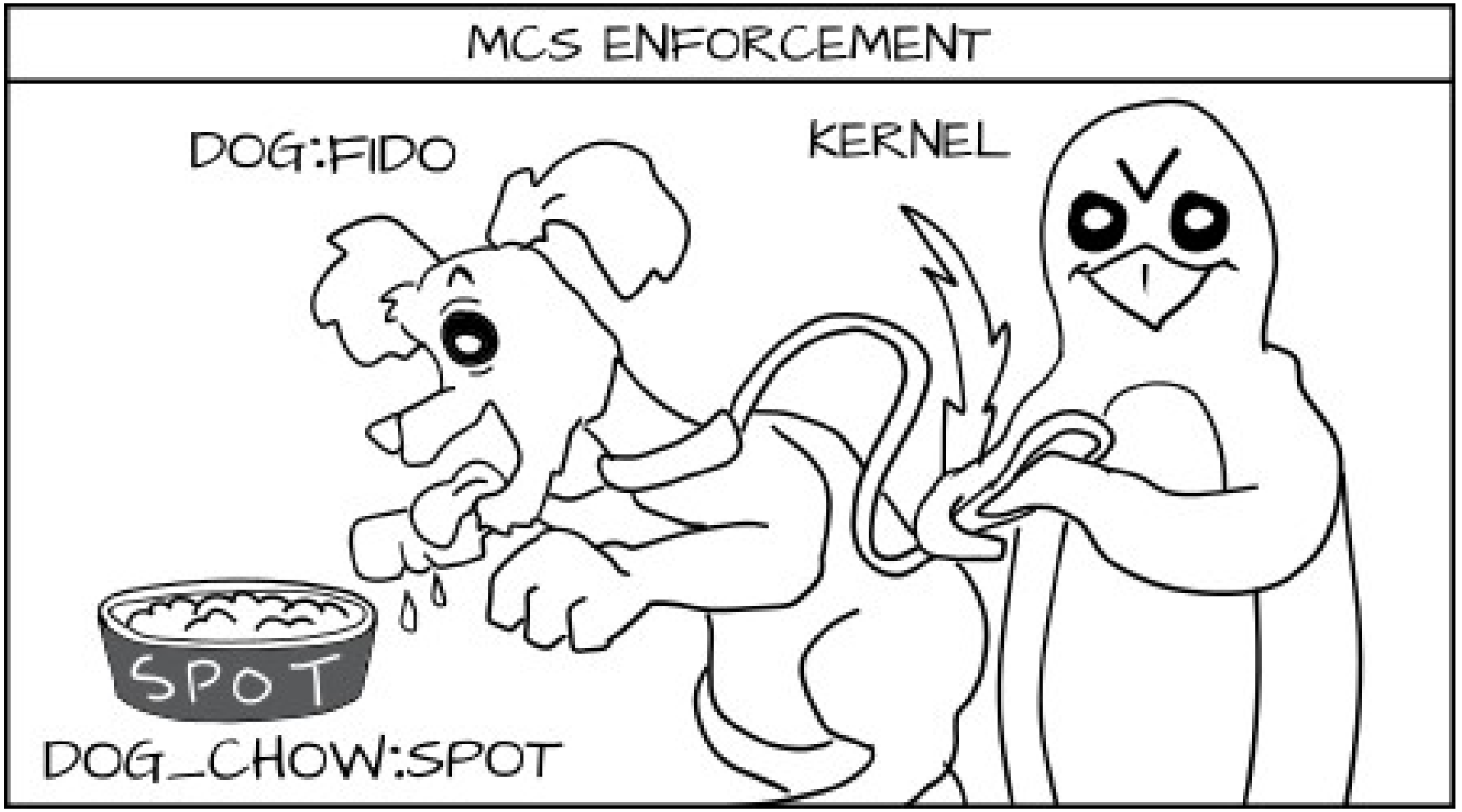
SELinux primer





Can same type of process be confined differently?

SELinux primer



Yes! With MCS enforcement!

SELinux primer

In real world...

2 processes: httpd_t

files under httpd: httpd_sys_content_t

So how to deny files from differ instances of httpd_t?

With MCS labels like s0:c1,c2 ; s0:c3,c4 etc

s0, s1, s2 – sensitivity levels

c1,c2,c3... - categories (up to 255)

So remember..

Every time you run setenforce 0, you make Dan Walsh weep

Dan is a nice guy and he certainly doesn't deserve that.

So what about other LSMs?

Feature	SELinux	AppArmor	grsecurity
Automated	No (audit2allow and system-config-selinux)	Yes (Yast wizard)	Yes (auto traning / gradm)
Powerful policy setup	Yes (very complex)	Yes	Yes
Default and recommended integration	CentOS / RedHat / Debian	Suse / OpenSuse	Any Linux distribution
Training and vendor support	Yes (Redhat)	Yes (Novell)	No (community forum and lists)
Recommend for	Advanced user	New / advanced user	New users
Feature	Pathname based system does not require labelling or relabelling filesystem	Attaches labels to all files, processes and objects	ACLs

http://www.cyberciti.biz/tips/selinux-vs-apparmor-vs-grsecurity.html

So what about other LSMs?

- AppArmor identifies file system objects by path name instead of inode

- There is no notion of multi-level security with AppArmor

- AppArmor user rather flat files based configuration

- SELinux supports the concept of a "remote policy server"

- There is no apparmor or grsec in android :)

Docker + SELinux

f20 policy: https://git.fedorahosted.org/cgit/selinux-policy.git/tree/docker.te?h=f20-contrib

What's there?

seinfo -t -x | grep docker

sesearch -A -s docker_t (and the rest)

or just unpack docker.pp with semodule_unpackage

Docker + SELinux

f20 policy: https://git.fedorahosted.org/cgit/selinux-policy.git/tree/docker.te?h=f20-contrib

What's there?

seinfo -t -x | grep docker

sesearch -A -s docker_t (and the rest)

or just unpack docker.pp with semodule_unpackage

How to use it?

man docker_selinux :)

Docker + SELinux

f20 policy: https://git.fedorahosted.org/cgit/selinux-policy.git/tree/docker.te?h=f20-contrib

What's there?

seinfo -t -x | grep docker

sesearch -A -s docker_t (and the rest)

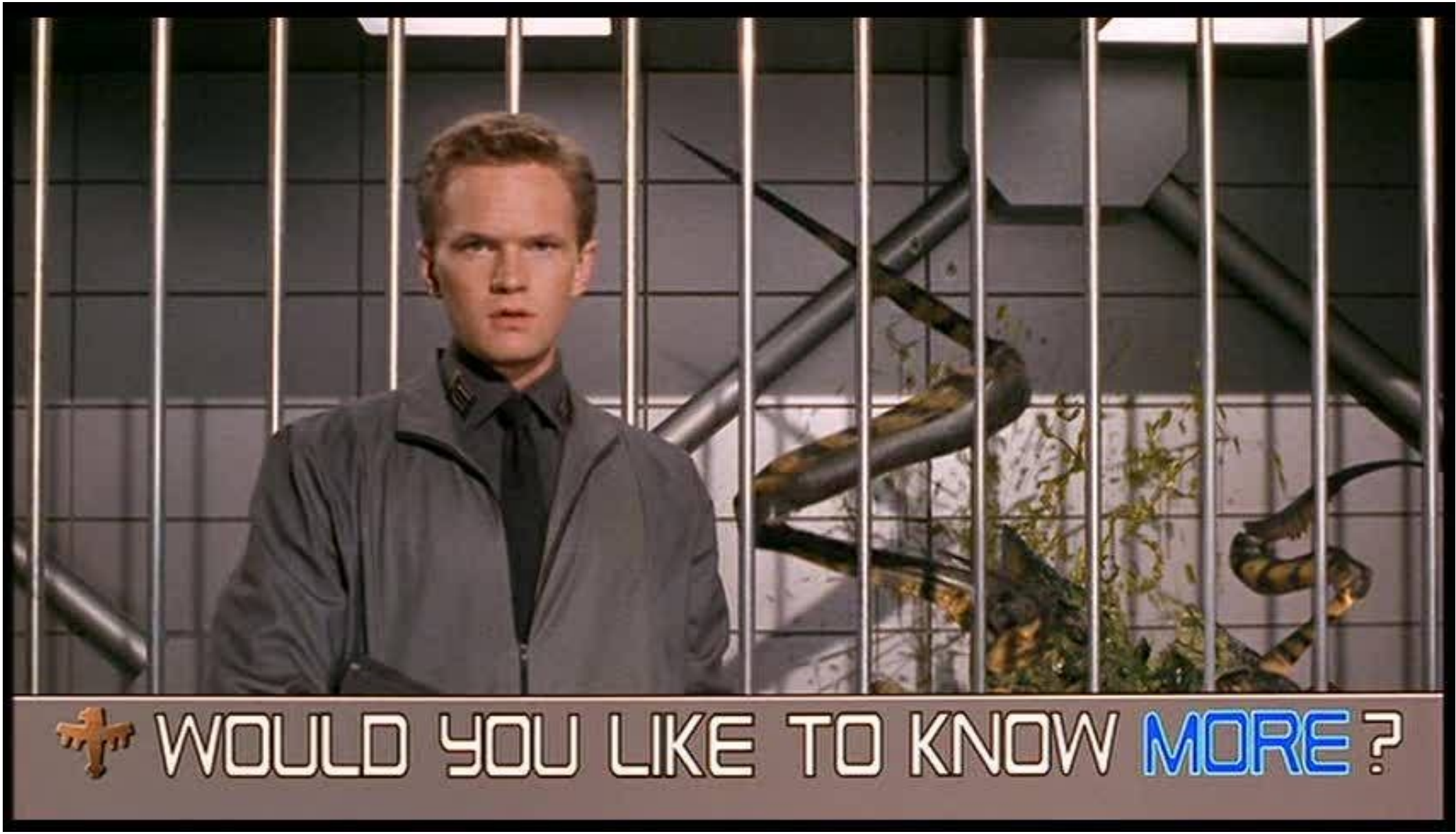
or just unpack docker.pp with semodule_unpackage

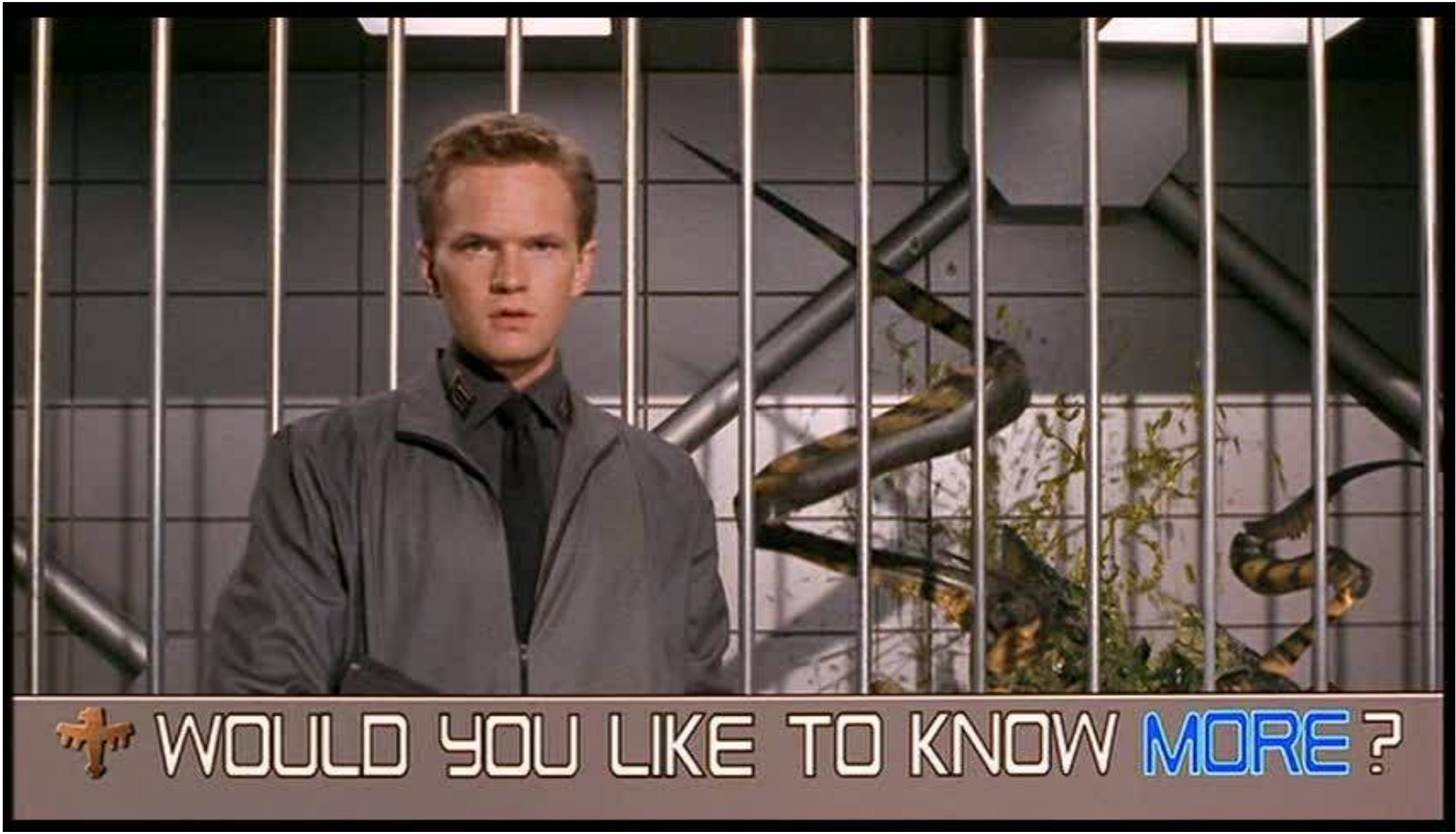
How to use it?

man docker_selinux :)

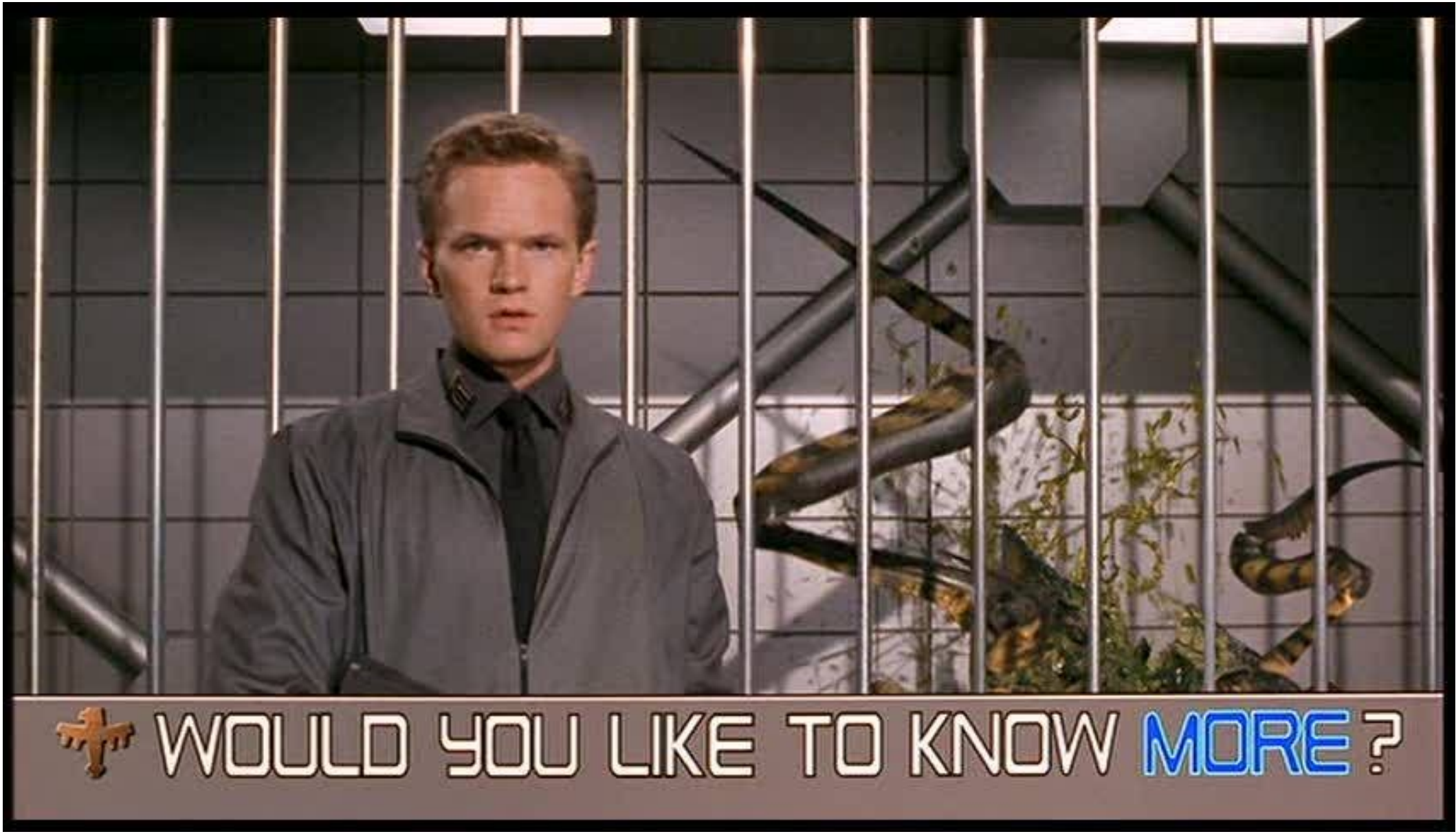
Permissive domains! semanage permissive -a docker_t

It's only in targeted policy (not for MCS)


WOULD YOU LIKE TO KNOW MORE?



stopdisablingselinux.com

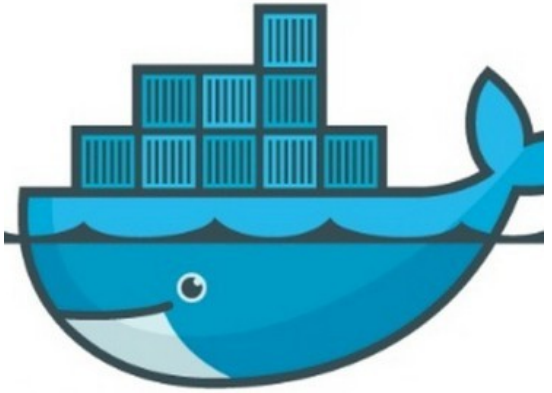


stopdisablingselinux.com

or...

Infosec meetup

Thank you :)



“Containers do not contain”

(orig. by Dan Walsh)

http://maciek.lasyk.info/sysop

maciek@lasyk.info

@docent-net