Security options for container implementations

Jay Coles
@container_doge

Who am I

http://doger.io
@container_doge

Triangle of Effort

Decreasing Skill Level

Increasing Effort

What they want

- Do not want to be detected
- Access to other customers information
- Access to other customers environments
- Adequate Storage/CPU/Mem/Network capacity
- Further ingress/infiltration on the network

How they do it

- Exploit an exposed service (does not need to have network access, eg in batch/queue processing)
- Pull down their toolset
- Start attacking the kernels
- Cement hold on system (command and control, process hiding)

What is security?

- Restrict access to other containers
- Prevent knowledge of other containers from leaking
- Ability to account for memory/cpu/network/disk usage
- Ability to control memory/cpu/network/disk resources
- Ability to detect and remove rouge processes

Usual Suspects

- Unix permissions
- Chroot
- Rlimit
- App Armor
- Selinux
- Capabilities
- Quotas
- Cgroups
- Seccomp
- ACLs

What does not work

- rlimits
- Quotas
- Blacklisting via ACLs

Capabilities

- CAP_SYS_MODULE
- CAP_SYS_RAWIO
- CAP_NET_BROADCAST
- CAP_MKNOD
- CAP_SYS_TTY_CONFIG
- CAP_AUDIT_WRITE
- CAP_AUDIT_CONTROL
- CAP_AUDIT_READ
- CAP_SYS_TIME
- CAP_MAC_OVERRIDE
- CAP_MAC_ADMIN
- CAP_NET_RAW
- CAP_SETPCAP
- CAP_SYSLOG
- CAP_WAKE_ALARM
- CAP_BLOCK_SUSPEND
- CAP_SYS_BOOT

Capabilities

- 'capsh' to drop capabilities
- Call instead of /sbin/init or entry point
- Have it invoke the init/entrypoint
- CAP_SETPCAP allows you to turn capabilties back on

cgroups

- Multiple protections in one
  - Accounting of resource usage
  - Limiting resource usage (cpu/mem)
  - Tracking of processes
  - Preventing/allowing device access

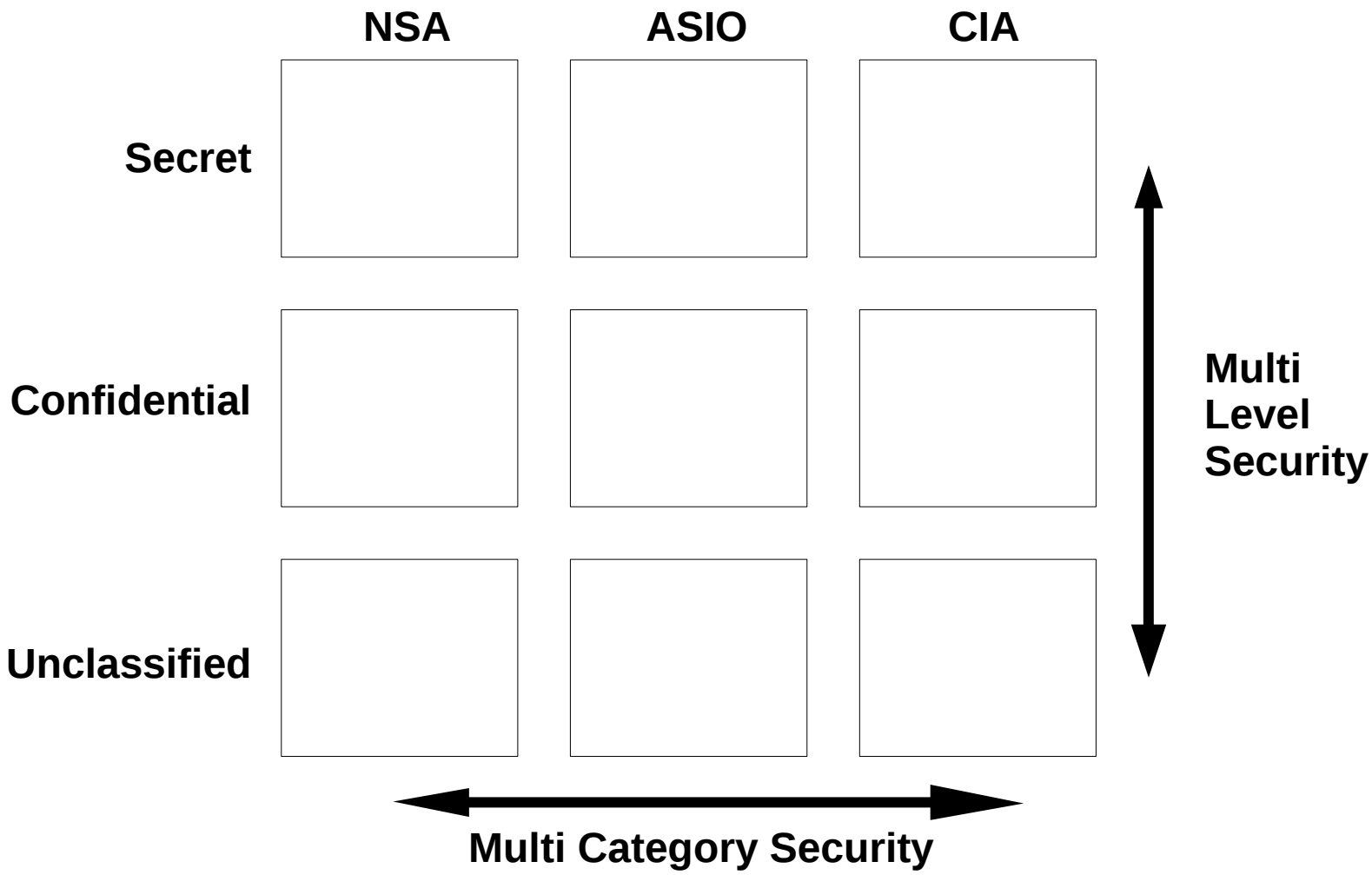
cgroups

```
dablitz@kodachi:~$ cd /sys/fs/cgroup/devices/

dablitz@kodachi:/sys/fs/cgroup/devices$ ls
cgroup.clone_children  devices.allow  notify_on_release
cgroup.procs           devices.deny   release_agent
cgroup.sane_behavior   devices.list   tasks

dablitz@kodachi:/sys/fs/cgroup/devices$ cat devices.list
a *:* rwm

dablitz@kodachi:/sys/fs/cgroup/devices$
```

App Armor vs selinux

Jay Coles
@container_doge

selinux


NSA
ASIO
CIA
Secret
Confidential
Unclassified
Multi
Level
Security
Multi Category Security

selinux

- 'runcon' is your friend
- 'chcon' to tag the files as belonging to a container
- Mainly going to be changing the security level
  - s0:c1,c4
- Will need appropriate policies/rules in place
  - This means a working selinux setup

seccomp

- Mount
- Acct
- Umount2
- Sethostname
- Swapon
- swapoff
- Reboot
- Adjtimeex
- Setdomainname
- init_module
- delete_module
- Quotactl
- finit_module
- Setns
- clock_adjtime
- kexec_load
- Nfsservct
- pivot_root
- pciconfig_iobase
- pciconfig_read
- pciconfig_write
- clock_settime
- Personality

Adding things in

- Can be patched in:
  - App Armor
  - Selinux
  - Capabilities
  - Cgroups
- Requires app support:
  - seccomp

Questions