Namespacing & Stacking the LSM

Casey Schaufler – Intel

John Johansen - Canonical
Linux Security Modules (LSM)

- Provide security
- Often MAC but not necessarily
- Kernel provides security
  - Hooks
  - Security field in various objects

- selinux, smack, apparmor, tomoyo, IMA/EVM, loadpin, yama
- proposed: LSMs: LandLock, CaitSith, Checmate, HardChroot, PTAGS, SimpleFlow, SafeName, WhiteEgret, shebang, S.A.R.A.
Namespacing the LSM

- Containers would like to be able to use alternate LSMs
- System container
  - lxd. run Ubuntu (apparmor) container on rhel (selinux) host
- Application confinement
  - snap using apparmor running on fedora (selinux base system)
  - Docker
  - flatpak
- ...
Stacking the LSM

- Have multiple LSMS enforcing at the same time
- minor LSM stacking merged in 4.2
- major LSM stacking iterating on networking issues
  - Move security blob maintenance into the infrastructure
  - Iterating on networking secids
What LSMs Want/Need

• Not every LSM has the same requirements
• System level confinement (confine the container)
  - eg selinux using MCS label per container
  - do NOT want either OR mediation
    • ie. selinux mediating tasks outside
    • container using different LSM not confined by selinux
• Application level confinement
  - Not every LSM supports
LSM Status

IMA
- Discussing namespaces
- Patch to ns IMA audit

Smack
- patches to label namespaces
- Per process rules

AppArmor
- Policy namespaces
- Virtualized interfaces
- Internal stacking

Audit
- wip to support namespaces
- Layering issues
Problems

- No consensus on what is needed
  - No agreement on what it means to be a container
  - Clarity on what is to be a namespace
  - LSMs need to namespace their attributes (xattrs...)
- secids
- Userspace interfaces
  - /proc/<pid>/attr/*
  - SO_PEERSEC
  - Secmark
- No security blob on namespaces
Problems cont.

- Subsystems not supporting namespaces (audit, ..)
- Missing mediation points around namespaces
- Hooks for LSM to update state when a task transitions NS
- cipso/calypso
- Seccomp no_new_privs
- Network layering issues