Per container system call filters using MOOL Kernel

Gayam Pradeep Kumar

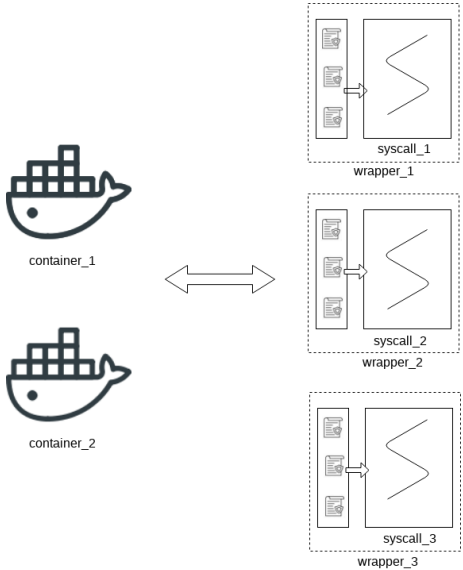
Guided: Prof. D. Janakiram

DOS Lab, IIT Madras

July 31, 2018

Table of Contents

Per container syscall filters



Per container syscall filters

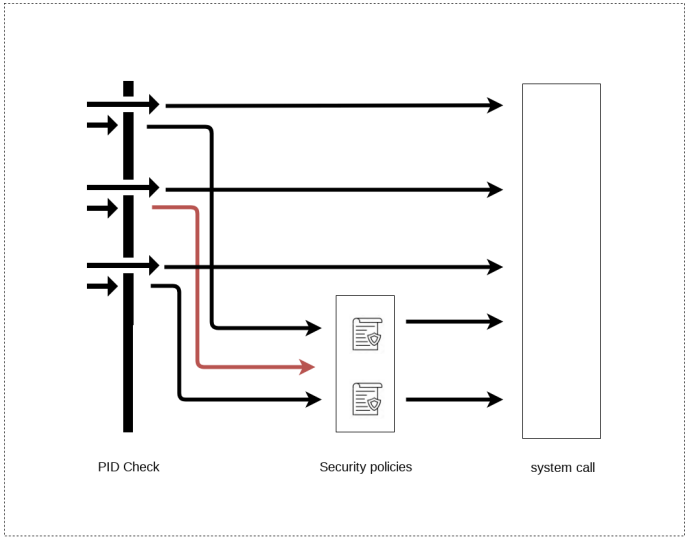


Table of Contents

Cgroups and Namespaces

- cpu
- memory
- devices
- net_cls
- blkio
- pids
- ...

- IPC
- Network
- Mount
- PID
- User
- UTS

Capabilities

- Capabilities are a way for running processes with some privileges, without having the need to grant them root privileges.
- They're are flags that tell the kernel what the application is allowed to do.
- Unprivileged containers are the new

Capability	Description
CAP_SET_PCAP	Modify process capabilities
CAP_SYS_MODULE	Insert/Remove kernel modules
CAP_SYS_RAW	IO Modify Kernel Memory
CAP_SYS_PACCT	Configure process accounting
CAP_SYS_NICE	Modify Priority of processes
CAP_SYS_RESOURCE	Override Resource Limits
CAP_SYS_TIME	Modify the system clock
CAP_SYS_TTY_CONFIG	Configure tty devices
CAP_AUDIT_WRITE	Write the audit log
CAP_AUDIT_CONTROL	Configure Audit Subsystem
CAP_MAC_OVERRIDE	Ignore Kernel MAC Policy
CAP_MAC_ADMIN	Configure MAC Configuration
CAP_SYSLOG	Modify Kernel printk behavior
CAP_NET_ADMIN	Configure the network
CAP_SYS_ADMIN	Catch all

Table: Capabilities

Linux Security Modules

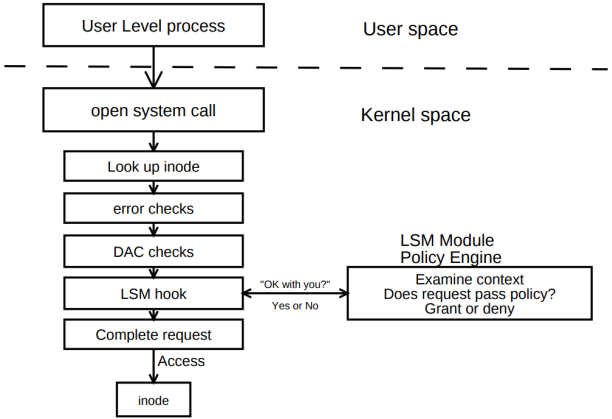


Figure: LSM Hook Architecture

LSM Hooks

- Task Hooks
- Program Loading Hooks
- File System Hooks
- IPC Hooks
- Module Hooks
- Network Hooks
- Other System Hooks

Table of Contents

Container security

- **Security in the context of Linux containers**

Container security

- **Security in the context of Linux containers**
  - All the above systems are used primarily to achieve isolation.

Container security

- **Security in the context of Linux containers**
  - All the above systems are used primarily to achieve isolation.
  - Namespace, cgroups, capabilities(unprivileged containers),LSM based MAC systems such as SELinux, AppArmor and seccomp-bpf.

Idea #1: As an alternative to LSM

- It's a kernel hardening method that helps to decide weather a system call should be allowed or not.

Idea #1: As an alternative to LSM

- It's a kernel hardening method that helps to decide weather a system call should be allowed or not.
- How does it fare against other security modules? How do we evaluate it?

Idea #2: To identify malicious containers