What is CAP_SYS_ADMIN?

CAP_SYS_ADMIN is a Linux capability that grants a huge, catch-all set of privileged operations, mounting filesystems, configuring namespaces, and much more, so broad it is often called "the new root". A container holding CAP_SYS_ADMIN can usually escape to the host, classically by abusing the cgroup `release_agent` mechanism or mounting host filesystems. It is sometimes added for convenience, but it effectively undoes capability dropping. Containers should run with it removed.

HOW IT WORKS

01 The escape and payload

A container with CAP_SYS_ADMIN has well-known escape routes:

- cgroup release_agent escape: mount the cgroup filesystem, set a `release_agent` script and `notify_on_release`, then trigger it so the host executes the attacker's script as root.
- Mounting host filesystems directly, then reading or writing host files (similar to a privileged container).

These turn the capability into root code execution on the host. Documented techniques shown for defenders.

> "THE NEW ROOT"
>
> *CAP_SYS_ADMIN is so broad it is treated as equivalent to root. Adding it to a container (often just for a mount or FUSE) usually hands an attacker a host escape via cgroups.*

HOW TO DEFEND

- Do not grant CAP_SYS_ADMIN to application containers; find a narrower capability or a different design.
- Drop all capabilities (`--cap-drop=ALL`) and add back only the specific, minimal ones required.
- Block added capabilities with Pod Security Standards (restricted) and admission control.
- Keep seccomp and AppArmor enabled to limit what even a capable container can do.
- Audit manifests for `SYS_ADMIN` in `capabilities.add` and test for cgroup-based escapes.

SOURCES

[1] Linux man-pages: capabilities(7)

[2] NIST SP 800-190 Application Container Security Guide

[3] MITRE ATT&CK: Containers Matrix

Find the container escape paths before an attacker does.

securelayer7.net/learn/containers/what-is-cap-sys-admin

Open online