What is a privileged container?

A privileged container is one started with Docker's `--privileged` flag (or a privileged Kubernetes security context), which gives it almost all Linux capabilities, access to host devices, and a relaxed seccomp/AppArmor profile. That effectively removes the isolation boundary: an attacker inside can mount the host disk and escape to the host in a few commands. It exists for legitimate niche workloads, but on a normal app it is a critical misconfiguration.

HOW IT WORKS

01 The escape and payload

From inside a privileged container, escaping to the host is short. A classic route is mounting the host disk:

- List host disks (visible because devices are exposed): `fdisk -l`
- Mount the host root filesystem: `mkdir /h && mount /dev/sda1 /h`
- Now read or write host files: drop a SSH key, a cron job, or `chroot /h` to operate as the host.

Another route abuses the cgroup `release_agent` to run a command on the host as root. Documented techniques shown for defenders.

> BARELY A BOUNDARY
>
> `--privileged` *is the single most impactful container misconfiguration: it hands the container host devices and capabilities, so a* container escape *is usually just a mount away.*

HOW TO DEFEND

- Never run application workloads as privileged. Treat `--privileged` as forbidden outside rare, audited infrastructure tools.
- Drop all capabilities and add back only the specific ones needed.
- In Kubernetes, block privileged pods with Pod Security Standards (restricted) and admission control.
- Run as non-root with a read-only root filesystem and seccomp on.
- Scan manifests for `privileged: true` and the `--privileged` flag in CI.

SOURCES

[1] Docker docs: Container runtime options

[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-a-privileged-container`

Open online