Contribution ID: **118** Type: **not specified**

Securing Container Runtimes with openat2 and libpathrs

Tuesday, 10 September 2019 18:00 (30 minutes)

Userspace has (for a long time) needed a mechanism to restrict path resolution. Obvious examples are those of FTP servers, Web Servers, archiving utilities, and now container runtimes. While the fundamental issue with privileged container runtimes opening paths within an untrusted rootfs was known about for many years, the recent CVEs (CVE-2018-15664 and CVE-2019-10152 being the most recent) to that effect has brought more light to the issue.

This is an update on the work briefly discussed during LPC 2018, complete with redesigned patches and a new userspace library that will allow for backwards-compatibility on older kernels that don't have openat2(2) support. In addition, the patchset now has new semantics for "magic links" (nd_jump_link-style "symlinks") that will protect against several file descriptor re-opening attacks (such as CVE-2016-9962 and CVE-2019-5736) that have affected all sorts of container runtimes and other programs. It also provides the ability for userspace to further restrict the re-opening capabilities of O_PATH descriptors.

In order to facilitate easier (safe) use of this interface, a new userspace library (libpathrs) has been developed which makes use of the new openat2(2) interfaces while also having userspace emulation of openat2(RESOLVE_IN_ROOT) for older kernels. The long-term goal is to switch the vast majority of userspace programs that deal with potentially-untrusted directory trees to use libpathrs and thus avoid all of these potential attacks.

The important parts of this work (and its upstream status) will be outlined and then discussion will open up on what outstanding issues might remain.

I agree to abide by the anti-harassment policy

Yes

I confirm that I am already registered for LPC 2019

Primary author: Mr SARAI, Aleksa (SUSE LLC)

Presenter: Mr SARAI, Aleksa (SUSE LLC)

Session Classification: Containers and Checkpoint/Restore MC