Contribution ID: **86** Type: **not specified**

systemd cgroup delegation and control processes

Monday, 12 September 2022 15:00 (20 minutes)

systemd manages the cgroup hierarchy from the root.
This is considered an exclusive operation and it is sufficient when system units don't encompass any internal cgroup structure.
To facilitate arbitrary needs of units, it is possible to delegate the subtree to the unit (a necessity for such units executing as unprivileged users).
However, the unified cgroup hierarchy comes with so called internal node constraint that prevents hosting processes in internal nodes of the cgroup tree (when controllers are enabled).

This creates a potential conflict between processes of the delegated unit and processes that systemd needs to run on behalf of the unit (e.g. ExecReload=).
Currently, it is avoided by putting systemd control processes into an auxiliary child cgroup directly under delegated subtree root.
This approach is broken when the subtree delegation is used to enable threaded cgroups since those require explicit setup and the auxiliary cgroup would miss that.
Generally, this is a problem of placing the control and payload processes within the cgroup hierarchy.

I'm putting forward a few patches that allow per-unit configuration of target cgroup of control and payload processes for units that have delegated subtrees.
This is a generic approach that keeps a backwards compatible default, avoids creation of unnecessary wrap cgroups and additionally allows new customization of control process execution.

It is a simple idea to present, this brings the topic up for discussion and comparison with similar situations that are affected by the internal node constraint too (e.g. joining a container) and the goal is to come up with a consent or at least the direction how to structure cgroup trees for delegated units that work well both for controller and threaded delegation.

This presentation and discussion will fit in a slot of 20 minutes.

I agree to abide by the anti-harassment policy

Yes

Primary author: KOUTNÝ, Michal

Presenter: KOUTNÝ, Michal

Session Classification: Service Management and systemd MC

Track Classification: LPC Microconference: Service Management and systemd MC