

cgroup v2 status update



Tejun Heo, facebook
kernel recipes 2016

timeline

- Preliminary work started with gradual cleanups in Sep, 2012.
- Experimental v2 unified hierarchy support implemented in Apr, 2014.
- cgroup v2 interface exposed in Nov, 2016 (Linux v4.5).
- memory, io and pids controller are supported currently.
- Unfortunately, cpu controller support is still out of tree.

what's v2 about?

Comprehensive and consistent hierarchical accounting and control of all significant resource consumptions in the system.

what does v2 do?

- Establish common resource domain across different resource types so that they can cooperate.
 - page cache writeback
 - resources spent for operations which can't be charged immediately
- Differing levels of granularity.
- Well-defined and consistent design in the core and across all controllers.
 - delegation
 - common resource models and interface conventions
- Other cleanups.
 - notification

cpu controller controversy

- cpu controller support is still out of tree.
- Disagreements around v2 core design features.
 - no-internal-processes constraint
 - process granularity, rgroup
- Discussion is still on-going.
 - <http://lwn.net/Articles/697366/>
 - <http://lkml.kernel.org/r/20160805170752.GK2542@mtj.duckdns.org>

other things

- pids controller added on v4.3.
- Namespace support added on v4.6. Still shaking out some bugs.
- rdma controller being worked on.

userland support

- systemd 232 will have full cgroup v2 support including cpu controller.
- It can also use cgroup v2 for process management while using v1 for resource control.
- Working on libvirt support.
- We are deploying cgroup v2 in combination with mainline systemd.

questions / comments?

Merci Beaucoup