Per-task I/O boost tracking

Monday, 20 September 2021 09:45 (30 minutes)

Sugov implements a rather simplistic concept of boosting I/O-bound tasks, through tracking I/O wakeups reported on each CPU and adjusting a synthetic boost value to potentially influence upcoming frequency changes. The actual boost value depends on a number of different conditions, like timings of the task wake-ups and CPU frequency update requests or the CPUfreq policy.
This makes things rather fuzzy and exposes the following drawbacks which might result in unexpected lost of I/O boost build-ups or undesired CPU frequency spikes:

1) Sugov does not differentiate between I/O boost request sources so it can't detect multiple unrelated tasks that do have sporadic I/O wake-ups.

2) As the boost value is being maintained per CPU, boost accumulated on one CPU might be lost upon task migration.

3) There is no guarantee that the task(s) that did trigger the I/O boost is/are still runnable on that CPU.

4) Relevant task uclamp restrictions are not being taken into account.

5) No notion of dependency on the actual device's performance and throughput. I.e. boosting CPU frequency might turn out to be pointless in case the device cannot cope with the increasing IO request rate.

This presentation shows how these shortcomings could be solved or at least mitigated by moving from per-core to per-task I/O boost tracking implementation.

I agree to abide by the anti-harassment policy

I agree

Primary authors: Ms MICHALSKA, Beata (Arm); EGGEMANN, Dietmar

Presenters: Ms MICHALSKA, Beata (Arm); EGGEMANN, Dietmar

Session Classification: Scheduler MC

Track Classification: Scheduler MC