



Contribution ID: 283

Type: **not specified**

Sharing PMU counters across compatible perf events

Monday, 9 September 2019 10:44 (22 minutes)

Hardware PMU counters are limited resources. When there are more perf events than the available hardware counters, it is necessary to use time multiplexing, and the perf events could not run 100% of time.

On the other hand, different perf events may measure the same metric, e.g., instructions. We call these perf events “compatible perf events”. Technically, one hardware counter could serve multiple compatible events at the same time. However, current perf implementation doesn’t allow compatible events to share hardware counters.

There are efforts to enable sharing of compatible perf events. To the best of our knowledge, the latest attempt was <https://lkml.org/lkml/2019/2/26/823>. Unfortunately, we haven’t make much progress on this front.

At Facebook we are investing on user space sharing of compatible performance counters to reduce the need for time multiplexing and the cost of context switch when monitoring the same events in several threads and cgroups. A kernel solution would be preferable.

In the Tracing MC, we would like to discuss how we can enable PMU sharing compatible perf events. This topic may open other discussions in perf subsystem. We think this would be a fun section.

I agree to abide by the anti-harassment policy

Yes

Primary authors: LIU, Song; CARRILLO CISNEROS, David (Facebook)

Presenters: LIU, Song; CARRILLO CISNEROS, David (Facebook)

Session Classification: Tracing MC