

Runtime power sharing among CPUs, GPUs and others

Thursday, 15 November 2018 11:20 (25 minutes)

Due to high performance demands systems tend to be over-provisioned, where it is not possible to run at peak power of each component. Even if each component has capability to report power and set power limits, there is no kernel level framework to achieve that. IPA addresses part of it, but on the systems in question thermal limits usually are not a problem, but sudden power overdraw is a bigger issue (particularly on unlocked systems). In addition, without proper power balance among components, they can starve each other. For example, in Intel KabyLake-G there are 4 big power consumers: CPUs, two GPUs and memory. If CPUs take most of the power, it will hurt graphics performance as the GPU will not be able to handle requests timely. So the power has to be managed at run time based on the workload demand.

I agree to abide by the anti-harassment policy

Primary author: PANDRUVADA, Srinivas (Intel)

Presenter: PANDRUVADA, Srinivas (Intel)

Session Classification: Power Management and Energy-awareness MC