



Contribution ID: 182

Type: **not specified**

Towards improved selection of CPU idle states

Thursday, 15 November 2018 09:50 (20 minutes)

The venerable menu governor does some things that are quite questionable in my view. First, it includes timer wakeups in the pattern detection data and mixes them up with wakeups from other sources which in some cases causes it to expect what essentially would be a timer wakeup in a time frame in which no timer wakeups are possible (because it knows the time until the next timer event and that is later than the expected wakeup time). Second, it uses the extra exit latency limit based on the predicted idle duration and depending on the number of tasks waiting on I/O, even though those tasks may run on a different CPU when they are woken up. Moreover, the time ranges used by it for the sleep length correction factors are not correlated to the list of available idle states in any way whatever and different correction factors are used depending on whether or not there are tasks waiting on I/O, which again doesn't imply anything in particular.

A major rework of the menu governor would be required to address these issues and it is likely that the performance of some workloads would suffer from that. That raises the question of whether or not to try to improve the menu governor or to introduce an entirely new one to replace it, or to do both these things simultaneously.

I agree to abide by the anti-harassment policy

Primary author: WYSOCKI, Rafael (Intel Open Source Technology Center)

Presenter: WYSOCKI, Rafael (Intel Open Source Technology Center)

Session Classification: Power Management and Energy-awareness MC