Integration of PM-runtime with System-wide Power Management

Tuesday, 10 September 2019 12:00 (45 minutes)

There are two flavors of power management supported by the Linux kernel: system-wide PM based on transitions of the entire system into sleep states and working-state PM focused on controlling individual components when the system as a whole is working. PM-runtime is part of working-state PM concerned about the opportunity to put devices into low-power states when they are not in use.

Since both PM-runtime and system-wide PM act on devices in a similar way (that is, they both put devices into low-power states and possibly enable them to generate wakeup signals), optimizations related to the handling of already suspended devices can be made, at least in principle. In particular:

* It should be possible to avoid resuming devices already suspended by runtime PM during system-wide PM transitions to sleep states. It should be possible to leave devices suspended during system-wide PM transitions to sleep states in PM-runtime suspend while resuming the system from those states.
* It should be possible to re-use PM-runtime callbacks in device drivers for the handling of system-wide PM.

These optimizations are done by some drivers, but making them work in general turns out to be a hard problem. They are achieved in different ways by different drivers and some of them are in effect only in specific platform configurations. Moreover, there are no general guidelines or recipes that driver writers can follow in order to arrange for these optimizations to take place. In an attempt to start a discussion on approaching this problem space more consistently, I will give an overview of it, describe the solutions proposed and used so far and suggest some changes that may help to improve the situation.