

Enhancements and adjustments of the thermal control subsystem

Rafael J. Wysocki

Intel

September 19, 2024

Terminology

Thermal zone : A thermal sensor (or equivalent) with a bundle of control and monitoring mechanisms (trip points, cooling devices and a thermal governor)

Zone temperature : Temperature value retrieved from the sensor

Thermal mitigation : Actions carried out to decrease the zone temperature

Trip point : Represents two levels of zone temperature triggering some kind of thermal mitigation (**trip temperature**) or the end of it (**trip low temperature**)

Trip hysteresis : Difference between trip temperature and trip low temperature

Cooling device : A mechanism pulling the zone temperature down

Cooling state : A level of cooling device activity

Thermal instance : A (parametrized) connection between a trip point and a cooling device

Thermal governor : Represents a thermal mitigation algorithm associated with a given thermal zone

Put lists of thermal instances into trip point descriptors

Summary

- Lists of thermal instances are located in thermal zone device objects.
- However, they represent connections between trip points and cooling devices.
- It is better to put them into trip point descriptors.

Patches

- <https://lore.kernel.org/linux-pm/2196792.irdbgypaU6@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/2954063.e9J7NaK4W3@rjwyssocki.net/>

Use sorted lists of trips for trip crossing detection

Summary

- Table of trip point descriptors (may not be sorted).
- When the thermal zone temperature changes, all trips need to be checked.
- Sorted lists are used anyway for notifications.
- Let each trip be always present in a sorted list.

Patches

- <https://lore.kernel.org/linux-pm/3316488.44csPzL39Z@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/10527854.nUPlyArG6x@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/1978487.PYKUYFuaPT@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/7708886.EvYhyI6sBW@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/3584535.iIbC2pHGD1@rjwyssocki.net/>

Define and use a thermal zone guard (for locking)

Summary

- Define a guard for thermal zone locking and use it in the thermal core.

Patches

- <https://lore.kernel.org/linux-pm/3241904.5fSG56mABF@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/4613601.LvFx2qVVIh@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/4940614.GXAFRqVoOG@rjwyssocki.net/>

Thermal zone initialization and exit protection

Summary

- User space may access new thermal zones before they are ready.
- Analogously, during thermal zone exit.
- There is a race between system suspend and thermal zone setup.

Patches

- <https://lore.kernel.org/linux-pm/2215885.irdbgypaU6@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/2973309.e9J7NaK4W3@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/3335807.44csPzL39Z@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/10548633.nUPlyArG6x@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/1997536.PYKUYFuaPT@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/7729094.EvYhyI6sBW@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/3603909.iIbC2pHGD1@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/3261209.5fSG56mABF@rjwyssocki.net/>

Use guards for managing locks in the thermal core

Summary

- Use guards for `thermal_list_lock`, `thermal_governor_lock` and the cooling device lock management.

Patches

- <https://lore.kernel.org/linux-pm/1822468.VLH7GnMWUR@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/2018087.usQuhbGJ8B@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/1921484.CQ0ukoFCf9@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/2641944.Lt9SDvczpP@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/2224279.Mh6RI2rZIC@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/3791590.MHq7AAxBmi@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/1890654.atdPhlSkOF@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/2495577.jE0xQCEvom@rjwyssocki.net/>
- <https://lore.kernel.org/linux-pm/863177860.0ifERbkFSE@rjwyssocki.net/>

Update cooling devices after unbinding them from thermal zones

Summary

- In case a thermal zone going away is driving the state of a cooling device.

Patches

- <https://lore.kernel.org/linux-pm/2227092.Icojqenx9y@rjwysocki.net/>

Other tweaks and testing enhancements

Other possible tweaks

- Let the Step-wise governor start to back off earlier.
- Use sorted lists of trip points in thermal governors.

Testing enhancements

- Circular buffer of notification messages.
- Mock cooling devices (including binding to trip points).
- `.set_trips()` testing support.