Towards Improved Selection of CPU Idle States

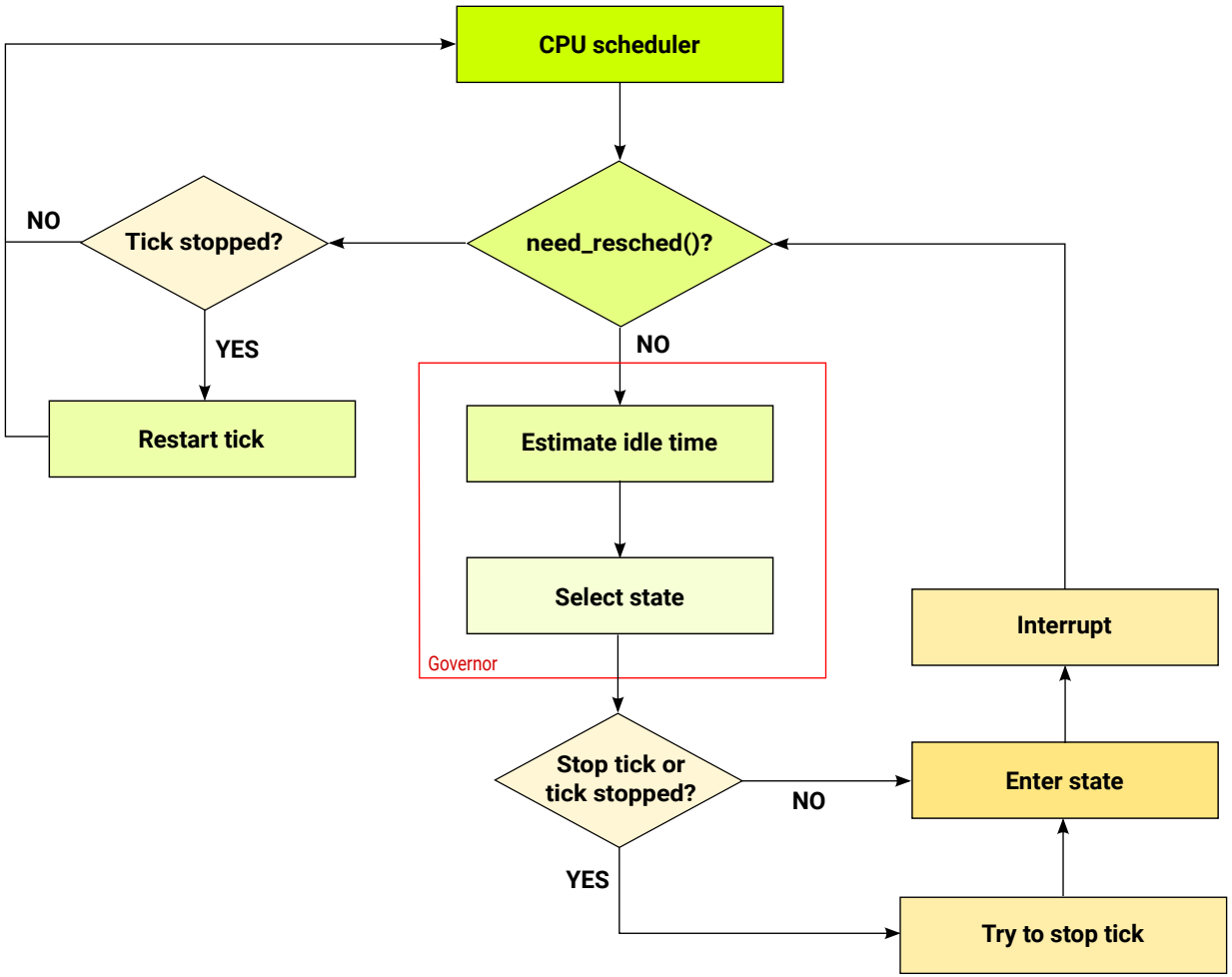
Rafael J. Wysocki

Intel Open Source Technology Center

November 15, 2018

Intel OpenSource TECHNOLOGY CENTER

CPU Idle Loop (Since Linux* 4.17)



Intel OpenSource TECHNOLOGY CENTER

* Other names and brands may be claimed as the property of others

CPU Idle States

Array of idle state objects per logical CPU

1. Ordered (shallowest first).
2. State parameters

   Target residency : The CPU shouldn't be interrupted for at least this time after asking the hardware to enter this idle state.

   Exit latency : Worst-case time span between asking the hardware to enter this idle state and the first instruction that can be executed after wakeup.

Intel OpenSource TECHNOLOGY CENTER

Idle State Selection Problem

Input

- Time till the closest timer (known exactly).
- Latency constraints (known exactly).
- Idle duration data from the past (distribution).
- Interrupt timings in the past (distribution)???

Questions

1. Is there a reason to believe that the CPU will be woken up before the closest timer?
2. What is the most likely time frame of that wakeup?

Intel OpenSource TECHNOLOGY CENTER

Issues In The menu Governor

Known issues in menu (Linux* 4.20-rc1 and later)

1. Timer wakeups included in pattern detection (correctness).
2. Pattern detection takes irrelevant data into account (overhead).
3. Extra latency constraints based on "I/O waiters" data (correctness).
4. Correction factors depending on "I/O waiters" data (correctness).

Observations

- Selecting idle states that are too shallow is bad for energy-efficiency.
- Selecting idle states that are too deep is bad for energy-efficienty and latency (performance).

Intel OpenSource TECHNOLOGY CENTER

* Other names and brands may be claimed as the property of others

Possible Mitigation Strategies

1. Introduce a new governor, leave `menu` alone (viable).
   - Need to fix switching governors at run time.
2. Fix `menu` (but it will become a different governor then).
3. Fix `menu` somewhat and introduce a new governor.

Intel OpenSource TECHNOLOGY CENTER

Opinions? Comments? Recipes? Suggestions?

Intel OpenSource TECHNOLOGY CENTER

Disclaimer

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at `www.intel.com`.

Intel OpenSource TECHNOLOGY CENTER