



**Linaro  
connect**

Hong Kong 2018

**OSPM 2018-04**

# **CPU Cluster Idling**

Ulf Hansson, Linaro

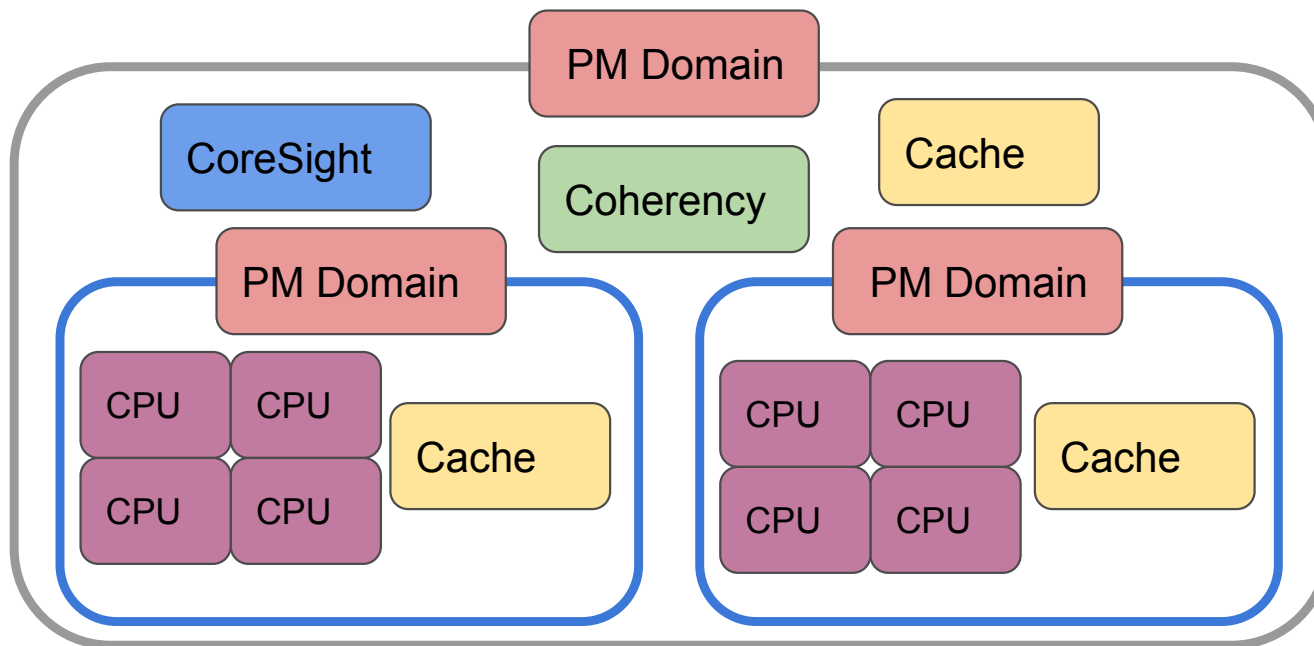


# CPU Cluster Idling

- Recap
- Status update
- Overview of current solution
- Limitations
- Open topics



# Recap



CPUIdle manages CPUs well, but does not scale for multi-cluster SMP systems and heterogeneous systems like big.LITTLE.

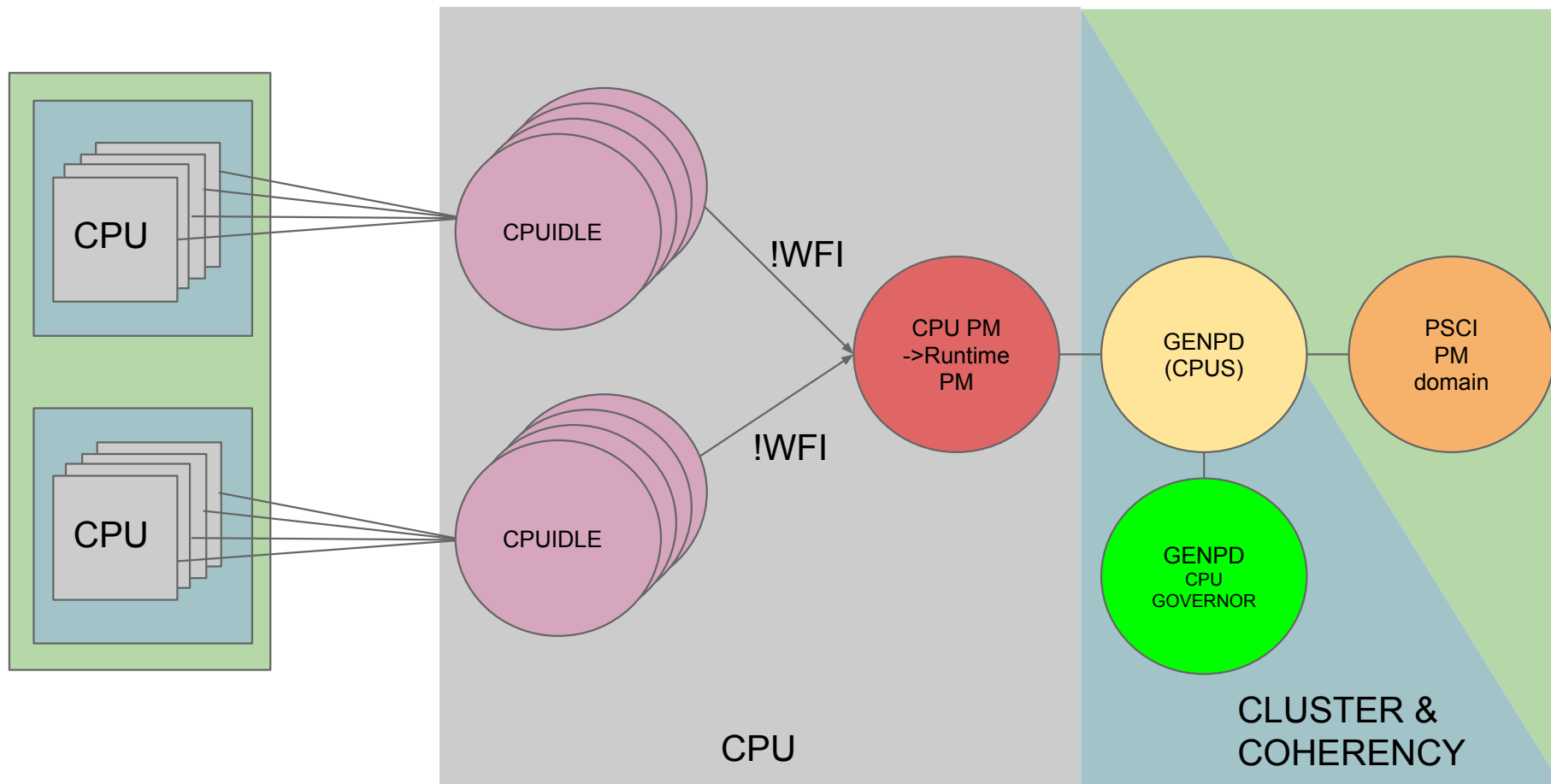


# Status update

- New version(s) have been posted.
  - <https://lwn.net/Articles/751650/> (v7)
  - [git.linaro.org/people/ulf.hansson/linux-pm.git](https://git.linaro.org/people/ulf.hansson/linux-pm.git) next
- Validated on 410c using ftrace and via collecting idle statistics.
- Real power measurements
  - Dragonboard 410c - quad core, one cluster.
  - Dragonboard 820c - quad core, two clusters
  - Updates only needed to DTS.



# Overview



# Initialization

Step 1:

- PSCI driver detects OSI mode support in FW.

Step 2:

- ARM64 request PSCI to init the CPU PM domain topology.
- PSCI enables the OSI mode.

Step 3:

- DT CPUIdle init -> PSCI DT idle init: Attach CPU to its PM domain.



# Enter/exit idle

## Enter:

- CPUIdle gov: Select an idlestate for the CPU.
- cpuidle-arm: `cpu_pm_enter()` -> `pm_runtime_put_sync()`.
  - genpd: ->`runtime_suspend()` the CPU device.
  - genpd: ->`power_off()` the CPU's PM domain.
  - genpd: ->`power_off()` the CPU's master PM domain (cluster).
  - genpd: CPU gov: Selects an idlestate for the master PM domain.
  - PSCI PM domain: Stores idlestate data in a per CPU variable.
- cpuidle-arm: `psci_cpu_suspend_enter()`.

## Exit:

- cpuidle-arm: `cpu_pm_exit()` -> `pm_runtime_get_sync()`.
- ...



# CPU hotplug

Allow genpd to power of the CPU's master PM domain:

- CPU offline: PSCI detach the CPU from its PM domain.
- CPU online: PSCI re-attach the CPU to its PM domain.





# Questions?

**Zzzzzzzzzzz.....**



# Limitations

- genpd: CPU gov ->**tick\_nohz\_get\_next\_wakeup(cpu)**

```
{  
    struct clock_event_device *dev = per_cpu(tick_cpu_device.evtdev, cpu);  
    return dev->next_event;  
}
```
- CPUs with > 1 idle states (excluded WFI)
  - genpd: CPU gov: Don't select/knows the idlestate for the CPU.



# Open topics

- CPUidle governor
  - genpd: CPU gov: Let it select the idlestate of the CPU as well?
  - Sharing functions for heuristics for next wakeup prediction, etc.
  - Play with irq-prediction.
- Caches/interconnects/irqchips/etc sharing cluster PM domain.
  - Regular devices managed by runtime PM - ok!
  - “Passive” runtime PM devices -> parent/child/device-links/”master”-domains or what?
  - cpu\_pm notifiers - striving to replace them!?
- EAS related:
  - Dev PM QoS constraints for CPUs, should we care?
  - Other?





**Linaro**  
**connect**

Hong Kong 2018

**Thank You!**

OSMP 2018-04

Ulf Hansson, Linaro

