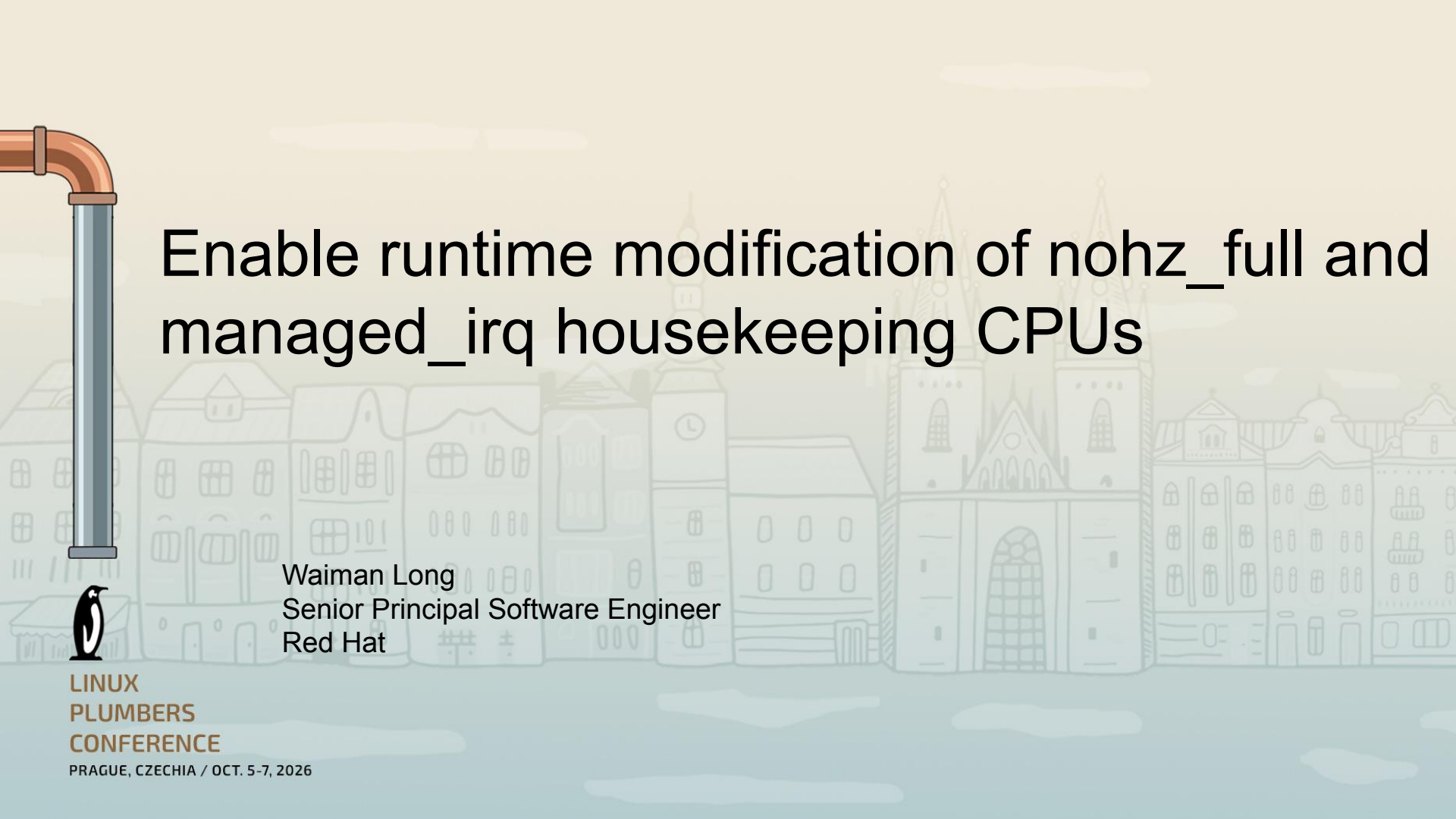




Enable runtime modification of nohz_full and managed_irq housekeeping CPUs

Waiman Long
Senior Principal Software Engineer
Red Hat



LINUX
PLUMBERS
CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026

Current CPU Isolation Support in the Kernel

CPU isolation can currently be done statically using the “isolcpus={domain,managed_irq}” and “nohz_full” boot parameters. They correspond to the following three housekeeping types:

- *HK_TYPE_DOMAIN*
- *HK_TYPE_MANAGED_IRQ*
- *HK_TYPE_KERNEL_NOISE*

It can also be done dynamically at runtime by using cpuset isolated partition feature which is currently limited to HK_TYPE_DOMAIN only. The goal is to enable runtime change to HK_TYPE_MANAGED_IRQ and HK_TYPE_KERNEL_NOISE using the same cpuset isolated partition feature.



LINUX

PLUMBERS CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026



HK_TYPE_KERNEL_NOISE Subtypes

HK_TYPE_KERNEL_NOISE used to be the combination of the following old housekeeping (HK) types:

- *HK_TYPE_TICK*
- *HK_TYPE_TIMER*
- *HK_TYPE_RCU*
- *HK_TYPE_KTHREAD*
- *HK_TYPE_WQ*
- *HK_TYPE_MISC*

Now HK_TYPE_WQ and HK_TYPE_KTHREAD have been merged into HK_TYPE_DOMAIN which is runtime changeable.



LINUX

PLUMBERS CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026



Making nohz_full/managed_irq Runtime Changeable

The current proposal is to make use of the existing CPU hotplug code to enable runtime changes to `HK_TYPE_KERNEL_NOISE` and `HK_TYPE_MANAGED_IRQ`.

- Shut down one-by-one the CPUs that have to be changed from non-isolated to isolated or vice versa
- Change the HK cpumasks
- Boot up the offline CPUs

Now the question is to what offline state should those CPUs be shut down to. The easiest way is to fully offline the CPUs (`CPUHP_OFFLINE`). The problem is that when a CPU goes down to `CPUHP_TEARDOWN_CPU`, all the other online CPUs will be stopped by the `stop_machine` mechanism until the offline is completed. If there are pre-isolated CPUs, this will cause latency spikes on the latency sensitive tasks they are working on.



LINUX

PLUMBERS CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026

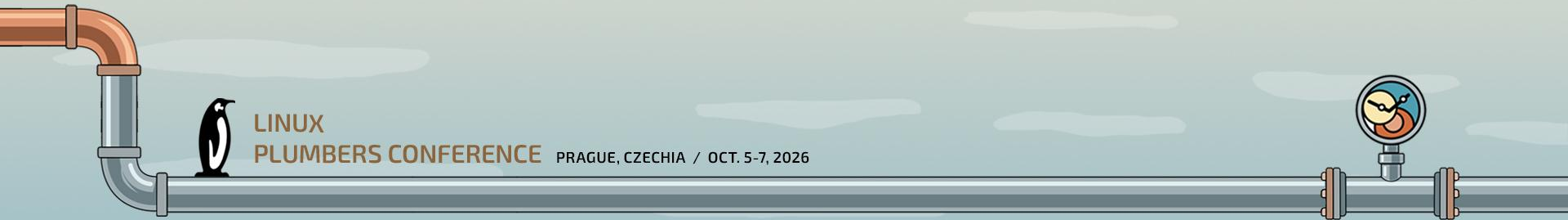


Hotplug State Shutdown Management

It is not problematic to fully shut down a CPU if there are no pre-isolated online CPU in the system. Otherwise, we can only shut down to CPUHP_AP_ONLINE_IDLE.

For HK_TYPE_TICK, HK_TYPE_TIMER and HK_TYPE_MANAGED_IRQ, it should be relatively easy to make all the necessary works be done by that hotplug state.

HK_TYPE_RCU, however, requires the CPU to be fully offline to work. Changing that will not be easy.



Hotplug State Shutdown Management (Cont)

The current proposal is to break `HK_TYPE_RCU` out from `HK_TYPE_KERNEL_NOISE` so that they can be managed independently. When there are no pre-isolated CPUs, the affected CPUs will be fully shut down. After modifying the HK cpumasks, those CPUs will be brought up again.

Once there are pre-isolated CPUs, `HK_TYPE_RCU` will stay untouched when we modified the `HK_TYPE_KERNEL_NOISE` and `HK_TYPE_MANAGED_IRQ` cpumasks. We can also provide an option to ignore the latency spikes and fully shut down the CPUs when necessary irrespective of the presence of pre-isolated CPUs.

Another alternative to avoid this latency spike problem is for users to use the “`rcu_nocbs`” boot parameter to create a set of no-callback CPUs and use only those CPUs in isolated partitions.



LINUX

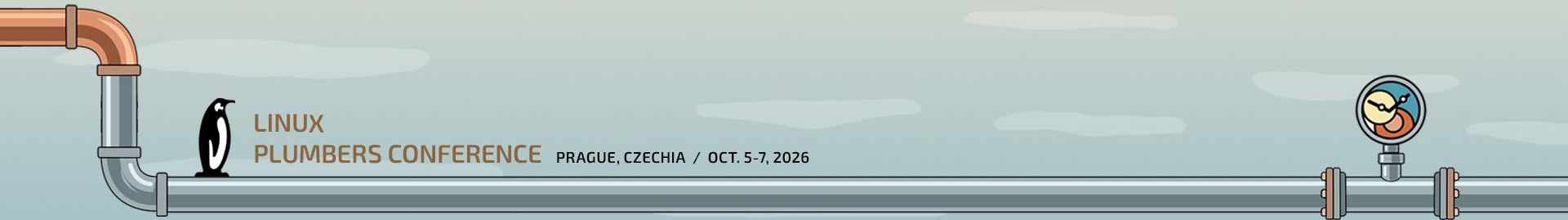
PLUMBERS CONFERENCE

PRAGUE, CZECHIA / OCT. 5-7, 2026

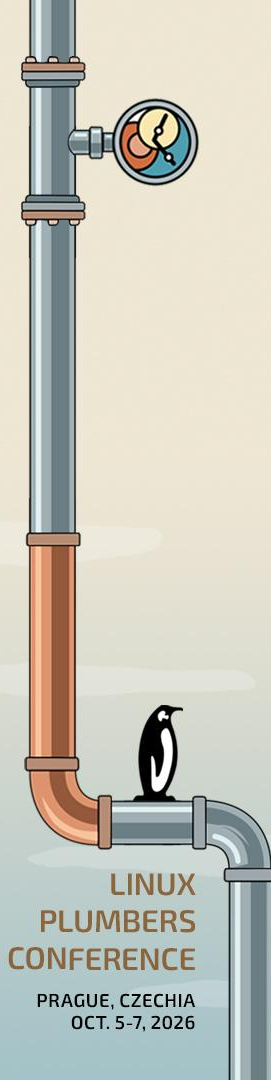


Future Works

1. There are still `HK_TYPE_MISC` and other ad hoc use of HK cpumasks at boot time to adjust behavior of various kernel subsystems which will not respond to runtime changes in HK cpumasks. Those subsystems will have to be modified to change their settings when the HK cpumasks are updated. Changes in HK cpumasks can be captured, for instance, by adding their own `cpuhotplug` notifiers to look for changes in the HK cpumasks.
2. Make RCU to change its `rcu_nocbs` cpumask without fully shutting down the CPU.



Q & A

A vertical pipe with a pressure gauge and a penguin on a horizontal section. The pipe is grey with orange-brown joints. A circular pressure gauge with a yellow face and blue needle is attached to the upper part. A penguin is perched on a horizontal section of the pipe near the bottom.

LINUX
PLUMBERS
CONFERENCE

PRAGUE, CZECHIA
OCT. 5-7, 2026