

Guest halt polling

The `cpuidle_haltpoll` driver, with the `haltpoll` governor, allows the guest vcpus to poll for a specified amount of time before halting.

This provides the following benefits to host side polling:

1. The `POLL` flag is set while polling is performed, which allows a remote vCPU to avoid sending an IPI (and the associated cost of handling the IPI) when performing a wakeup.
2. The VM-exit cost can be avoided.

The downside of guest side polling is that polling is performed even with other runnable tasks in the host.

The basic logic as follows: A global value, `guest_halt_poll_ns`, is configured by the user, indicating the maximum amount of time polling is allowed. This value is fixed.

Each vcpu has an adjustable `guest_halt_poll_ns` ("per-cpu `guest_halt_poll_ns`"), which is adjusted by the algorithm in response to events (explained below).

Module Parameters

The `haltpoll` governor has 5 tunable module parameters:

1. `guest_halt_poll_ns`:

Maximum amount of time, in nanoseconds, that polling is performed before halting.

Default: 200000

2. `guest_halt_poll_shrink`:

Division factor used to shrink per-cpu `guest_halt_poll_ns` when wakeup event occurs after the global `guest_halt_poll_ns`.

Default: 2

3. `guest_halt_poll_grow`:

Multiplication factor used to grow per-cpu `guest_halt_poll_ns` when event occurs after per-cpu `guest_halt_poll_ns` but before global `guest_halt_poll_ns`.

Default: 2

4. `guest_halt_poll_grow_start`:

The per-cpu `guest_halt_poll_ns` eventually reaches zero in case of an idle system. This value sets the initial per-cpu `guest_halt_poll_ns` when growing. This can be increased from 10000, to avoid misses during the initial growth stage:

10k, 20k, 40k, ... (example assumes `guest_halt_poll_grow=2`).

Default: 50000

5. `guest_halt_poll_allow_shrink`:

Bool parameter which allows shrinking. Set to N to avoid it (per-cpu `guest_halt_poll_ns` will remain high once achieves global `guest_halt_poll_ns` value).

Default: Y

The module parameters can be set from the debugfs files in:

```
/sys/module/haltpoll/parameters/
```

Further Notes

- Care should be taken when setting the `guest_halt_poll_ns` parameter as a large value has the potential to drive the cpu usage to 100% on a machine which would be almost entirely idle otherwise.