

Per Thread Queues (PTQ)

Wednesday, 26 August 2020 07:45 (45 minutes)

In this talk we introduce Per Thread Queues (PTQ). PTQ is a type of network packet steering that allows application threads to be assigned dedicated network queues for both transmit and receive. This facility provides highly granular traffic isolation between applications and can also help facilitate high performance when combined with other techniques such as busy polling. PTQ extends both XPS and aRFS.

A key concept of PTQ is “global queues”. These are a device independent, abstract representation of network queues. Global queues are as their name implies, they can be treated as managed resource across not only a system, but also across the a data center similar to how other resources are managed across a datacenter (memory, CPU, network priority, etc.). User defined semantics and QoS characteristics can be added to global queues. For instance, queue #233 in the datacenter might refer to a queue with QoS properties specific to handling video. Ultimately in the data path, a global queue is resolved to a real device queue that provides the semantics and QoS associated with the global queue. This resolution happens per a device specific mapping functions that maps a global queue to a device queue.

Threads may be assigned a global queue for both transmit and receive. The assignment comes from pools of transmit and receive queues configured in a cgroup. When a thread starts in a cgroup, the queue pools of the cgroup are consulted. If a queue pool is configured, the kernel assigns a queue to the thread (either a TX queue, RX queue, or both). The assigned queues are stored in the threads task structure. To transmit, the mapped device queue for the assigned transmit queue is used in lieu of XPS queue selection; for receive, the mapped device queue for the assigned receive queue is programmed into the device via `ndo_rx_flow_steer`.

This talk will cover the design, implementation, and configuration of PTQ. Additionally, we will present performance numbers and discuss some of the many ways that this work can be further enhanced.

I agree to abide by the anti-harassment policy

I agree

Primary author: HERBERT, Tom

Presenter: HERBERT, Tom

Session Classification: Networking and BPF Summit

Track Classification: Networking & BPF Summit