Contribution ID: **393** Type: **not specified**

Marking Packets With Rich Metadata

Friday 20 September 2024 15:30 (30 minutes)

Currently, the only way to attach a piece of information to an sk_buff that will travel with it through the network stack is the mark field.

Once set, the mark can be read in firewall rules, used to drive routing, and accessed by BPF programs, among other uses. This versatility leads to fierce competition over the mark's bits. Being just 32 bits wide, it often ends up limiting its practical applications.

Interestingly, there is already support for attaching more than just four bytes of metadata to a packet from the XDP context. In this presentation, we want to discuss how to extend this concept so that packet metadata can be accessed by other BPF programs which run later in the stack on the RX path, such as sk_lookup, reuseport, and socket filter.

Furthermore, we want to examine how packet metadata could be consumed by user-space programs using well-known patterns from the socket API, such as socket options and ancillary messages (cmsg).

During the talk, we would also like to highlight how attaching rich metadata to packets enables new and exciting applications such as:
* Tracing packets through layers of the network stack, even when crossing the kernel-user space barrier.
* Metadata-based packet redirection, routing, and socket steering with early packet classification in XDP.
* Extraction of information from encapsulation headers and passing it to user space, or vice versa.

We also want to explore how metadata could be structured to allow different users to share it without interference by leveraging the power of BTF based on prior work in that field.

Primary authors: FABRE, Arthur (Cloudflare); SITNICKI, Jakub (Cloudflare)

Presenters: FABRE, Arthur (Cloudflare); SITNICKI, Jakub (Cloudflare)

Session Classification: eBPF Track

Track Classification: eBPF Track