Linux Plumbers Conference 2025

Contribution ID: **402** Type: **not specified**

Challenges with implementing in-kernel FQDN policies using eBPF

Container networking plugins for Kubernetes like Cilium currently implement Fully Qualified Domain Name (FQDN) based DNS network policies using a user-space DNS proxy to intercept the DNS to IP mappings and plumb CIDR based policy into bpf maps.

This architecture introduces some challenges since any downtime with the the userspace proxy would result in DNS resolution failure for all workloads on the node. Solutions like introducing another userspace proxy in a high availability setup can improve the reliability, but it comes with the cost of introducing a protocol to communicate policy state and other metadata to the proxy. Thanks to some of the enhancements in eBPF it is now possible to implement a DNS parser natively in bpf. Cilium already enforces CIDR based policy in the kernel. Combining these two solutions we were able to get rid of the userspace proxy completely, eliminating the dataplane and control plane coupling and resulting in improved tail latencies.

This system uses a set of stream parser and stream verdict bpf programs to even support DNS over TCP. The path to implementing this however was not easy. Features like DNS compression make implementing such a system tricky, requiring us to understand some of the internals of verifier. In some scenarios, upgrading to newer kernel would simply resolve the issue. But not before spending days, if not weeks trying to reason about verifier's behavior. And then some more time bisecting to understand which commits fixed the issue.

This talk will dive into the details of how the system was built, share our experience during the development process and leave some room to discuss how we could improve the user experience (UX) for relatively new developers. The answer to these challenges could simply be documenting the behavior at the intersection of loops and other features or even a better abstraction in bpf that allows simplifying the work verifier needs to perform in these scenarios.

Primary author: MALLA, Hemanth (Microsoft)

Presenter: MALLA, Hemanth (Microsoft)

Session Classification: eBPF Track

Track Classification: eBPF Track