

Whats happening in Linux network development

Florian Westphal

4096R/AD5FF600 fw@strlen.de
80A9 20C5 B203 E069 F586
AE9F 7091 A8D9 AD5F F600

May 2017

HW Offloading

- Switching (switchdev framework)
- fib (router)
- IPsec
- TLS (tx-only rfc patch set only atm)
- tc offloading

XDP (eXpress data path)

- EBPF based
- JITed code running in kernel space at lowest level, i.e. in driver
- before any (skb) memory allocation
- use cases:
 - early drop
 - load balancers (packet rewrite and tx on same interface)
- generic xdp exists but it doesn't avoid skb alloc overhead

recent TCP changes

- BBR congestion control
- RACK (time-based loss detection)
- 1ms tcp timestamp clock resolution
- uptime concealment (per host random offsets)
- generic busypoll

- ipv6 segment routing
- optimization work:
 - UDP
 - MacSec
 - IPsec (lockless SA lookups)
 - conntrack timer removal
- nftables extensions: fib, hash
- wip
 - full (tcp, udp, ..) zerocopy via sendmsg MSG_ZEROCOPY
 - lockless qdisc
 - new AF_PACKET mode to expose rx/tx rings to userspace
 - ipv6 fib rwlock work