cls_bpf/eBPF updates since netdev 1.1

Daniel Borkmann
<daniel@iogearbox.net>
Noiro Networks / Cisco Systems

netdev 1.2, tc workshop
Tokyo, October 5, 2016
eBPF cls_bpf Programs

- Access to collect metadata tunnel key and options
  - vxlan, geneve, gre, ipip, ipip6, ip6ip6
- Direct packet access (read, write, helpers)
- Event output helper
- Cgroups v2 support
- Several other new BPF helper functions
  - Better checksum fixup support (l3, l4, csum_complete)
  - Improvements in header rewrites (f.e. helping NAT64)
  - skb metadata mangling (pkt_type, hash, etc)
- Hardware offloading (Netronome SmartNICs)!
- Various fixes, optimizations, misc improvements
eBPF Core

- Hash map preallocation
- New verifier argument types for helpers
  - Possibly uninitialized stack buffers (helper error paths adapted)
  - Possible NULL arguments for stack buffers
- Improvements in verifier search pruning
- RCU and refcounting fixes
- Several test_bpf and test_verifier additions
- Various other misc improvements
eBPF (arch/) JITs

- Available as of today: x86_64, arm64, ppc64, s390
- Generic constant blinding infrastructure
- ppc64
  - Initial eBPF JIT added!
  - Tail call support added
- arm64
  - Tail call support added, jiting of BPF_JSET
  - Various optimizations (register usage, stack frames)
  - Still missing BPF_XADD, otherwise complete
- s390
  - Recache fix related to skb->data/hlen
  - Fix to reduce maximum program size