bridge conntrack status update

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

Red Hat

July 2017
Plan/wishlist

- allow to track connections that pass through a bridge
  - no NAT support
  - no (automated) strip of headers (e.g. pppoe)
- deprecate/remove call-iptables infra
- remove skb->nf_bridge
status now

- created nf_conntrack_bridge.c
  - creates conntrack hook points at NFPROTO_BRIDGE
  - calls ipv4 or ipv6 conntrack depending on skb protocol type
  - appears to work, but ...
status now, the problems

ip fragmentation
  ▶ added defrag hooks
  ▶ no more nf_bridge info, placed max_frag_size in bridge cb again
  ▶ easy to handle refrag in bridge postrouting, but ...
  ▶ what to do if skb is for local machine?
  ▶ it could also be forwarded
  ▶ skb->cb is not preserved
problems, part 2

- bridge can clone skbs when forwarding, multiple skbs refer to same nf_conn entry
- breaks assumptions in conntrack for NEW packets wrt. conntrack extension area expansion
- already have this problem with call-iptables infra + nfqueue
- minor issues:
  - tracks all packets by default
  - some code duplication w. bridge netfilter
solutions

fragmentation

1. place `conntrack_in` hook into `forward` instead of `pre"routing"`
   - solves the "input problem" for fragmentation: we only see bridged packets
   - major issue: can’t use conntrack for bridge vs. local decisions anymore

2. refragment also in `input` too (but: yuck)