draft-perreault-softwire-lw4over6-pcp-00

IETF 87, Berlin
SOFTWIRE meeting
2013-07-30

Chongfeng Xie <xiechf@ctbri.com.cn>
Simon Perreault <simon.perreault@viagenie.ca>
Cathy Zhou <cathy.zhou@huawei.com>

Background

- Implemented both DHCP and PCP provisioning for lw4o6 (client, server, and DHCP relay)
 - PCP: a breeze
  - DHCP: hell

Problems with DHCP provisioning

- Needs a DHCPv4-over-IPv6 shim
  - NOT as easy as it seems
  - Required hacks to work around broadcast vs unicast
  - Being redone differently by DHC (they realized said hacks were really bad)
- Existing DHCP servers don't have logic to allocate port sets
  - Need to add a big chuck of fairly complex code on top of old DHCP server code that nobody wants to touch
- DHCP ties leases to individual addresses, not port sets
  - DHCP logic implemented everywhere needs to be ripped open deeply and rewritten
  - Needs to be done on clients, servers, relays, everywhere. Really really sucks.

Solutions with PCP provisioning

- PCP does all this already.

- Implementing PORT_SET on top of MAP is easy.

- No need to touch old code on which everyone depends for day-to-day life.

- When a DHCP server allocates a port set, it needs to configure the lwAFTR with that new allocation.

  - What protocol makes most sense for doing that?
    Third-party PCP.

  - If you end up with DHCP from the client to the DHCP server, then PCP to the lwAFTR, why not eliminate the middle-man and let the client speak PCP directly with the lwAFTR?

Lightweight 4over6 setup
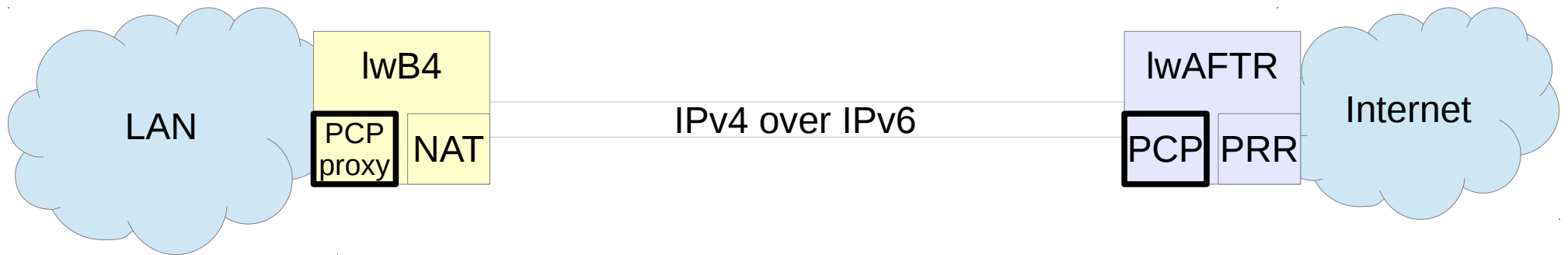
- NAT function in lwB4 is provisioned using PCP MAP requests with PORT_SET option.

  – PCP over IPv4, inside the tunnel.

- No use of DHCPv4.

- PCP server in lwAFTR controls PRR (port-range router) function

- PCP proxy in lwB4 controls NAT function

Provisioning items

- Tunnel
  - Outside IPv6 addresses:
    - Remote: DHCPv6 OPTION_AFTR_NAME
    - Local: Default Address Selection
      (exactly like draft-ietf-softwire-unified-cpe)
  - Inside IPv4 addresses:
    - Remote: 192.0.0.1
    - Local: 192.0.0.2
      (exactly like regular DS-Lite)
- NAT
  - PCP MAP request sent from 192.0.0.2 to 192.0.0.1
  - PORT_SET option with Size = 65535

PCP Proxy on lwB4

- PCP client on lwB4 maintains sufficient **pool** of external ports.

    - Sufficient for implicit and explicit mapping creation.

- PCP server on lwB4 picks from the pool to serve requests from clients on LAN.

- (Effect is like a PCP proxy, except that requests from LAN do not immediately trigger requests to lwAFTR.)

Failover mechanisms

- ICMP type 1 code 5 means from lwAFTR to lwB4 means that there is no matching entry in the lwAFTR binding table.
    - From draft-ietf-softwire-lw4over6
    - The corresponding external port set is deprovisioned
    - Triggers a new PCP request for a port set
- ANNOUNCE
    - Can be multicasted for efficient fast recovery of multiple lwB4's
- Open issue: Do we need just one mechanism (which one?) or both?

Next steps

- What are they?