

draft-ietf-shim6-reach- detect-01.txt

IETF64, shim6 wg, Vancouver, 2005-11-11

John Loughney (presenting)
Iljitsch van Beijnum (slides)

(un)reachability

- Assume list of potentially reachable local addresses of appropriate scope
- Addresses for correspondent and their preferences are known from shim context establishment/signalling

preferences

- Preference for each address determined by:
 - "Strong" value p1: higher is *always* better
 - "Weak" value p2: determines load balancing weight relative to other addresses with same p1 value
- Administrative interface undefined as yet

unidirectional reachability

- Relatively uncommon for normal failures
- **But:** assuming bidirectional reachability when it's unidirectional is **deadly**
- Ingress filtering can easily create unidirectional reachability
- So: work with unidirectional paths

keepalives

- Detect failures through keepalives
- Keepalive: IPv6 + 8 byte shim6 header
- Keepalives suppressed when:
 - No incoming data traffic (= context idle)
 - Already normal outgoing traffic
- So only keepalives for unidirectional protocols + some active/idle transitions

quick vs full

- Keepalive interval 10 seconds
- When *one* keepalive missed: quick reachability check with current address pair
- When quick check fails: full exploration
- Quick vs full: packet format largely the same, quick is shorter (fewer struct_p's)

packet format

- Encapsulated in shim6 header (allow for reuse in other protocols)
- Refer to both probes received from other side as well as probes we sent but haven't been acknowledged
- Probe references through "struct_p"
- Type/action bits include "request reply"

fields

- Housekeeping: type/action (8 bits), number of probes sent and received (both 8 bits), optional data length (8 bits)
- struct_p for this probe
- 0 or more struct_p for probes sent earlier
- 0 or more struct_p for probes received
- Options

struct_p

- Probe structure:
 - source locator (128 bits)
 - destination locator (128 bits)
 - sent timestamp (32 bits, in ms)
 - time since reception (32 bits, in ms)
 - nonce (32 bits)
 - seqnum (32 bits)

granularity

- Keepalives are ULID-to-ULID (per context)
- Address pair probing is host-to-host
- So contexts put address pairs in "to test" host-wide pool (until context idle 240 sec)
- Test pairs with highest p1 preference first
- When working address pair found: keep testing pairs with at least one better p1

timing

- Currently:
 - Limit probes to average of 60 per 300 seconds host-wide
 - But use aggressive timeouts (starting at 200 ms) first 15 - 20 seconds
- Is too conservative for busy hosts, though