Shim6 Protocol

draft-ietf-shim6-l3shim-00.txt
Erik Nordmark
erik.nordmark@sun.com

draft-ietf-shim6-functional-dec-00.txt
Marcelo Bagnulo

Outline

- Overview of approach
- Changes since last draft versions
- Open Issues
- Next Steps

Overview

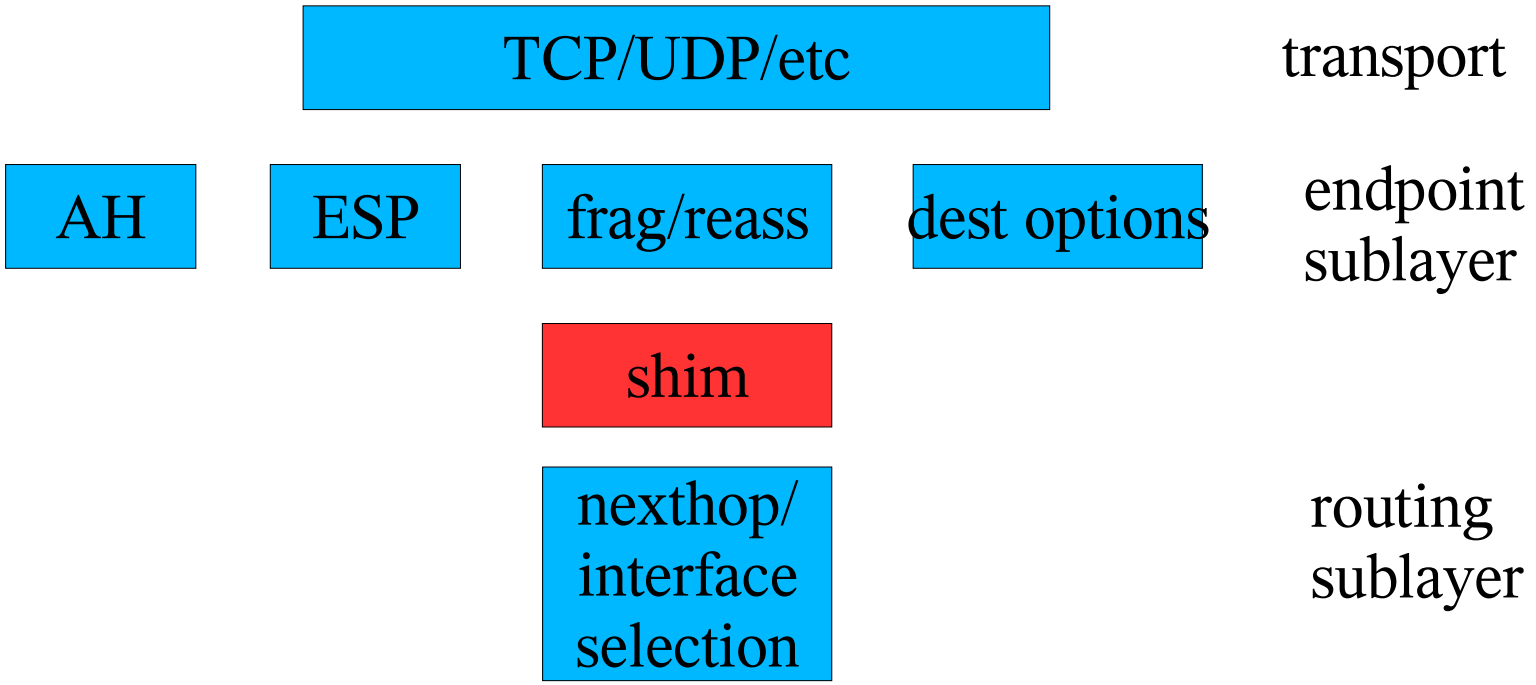
- No separate ID name space
- Placement of the L3 shim
- Assumptions about the DNS
- Deferred context establishment
- 4-way exchange for capability detection and context establishment

No Separate ID name space

- ULID – upper-layer ID
  - The 128-bit quantity which is used above the shim layer
  - Just one of the IPv6 addresses
- The set of locators (from AAAA records) are candidates for being the ULID
- The ULID is what's seen by TCP, applications etc
- Underneath the shim switches to use different locator(s) after a failure

Placement of the L3 shim

- Above the IP routing sublayer, below the IP endpoint sublayer
  - Below fragmentation, IPsec



Assumptions about the DNS

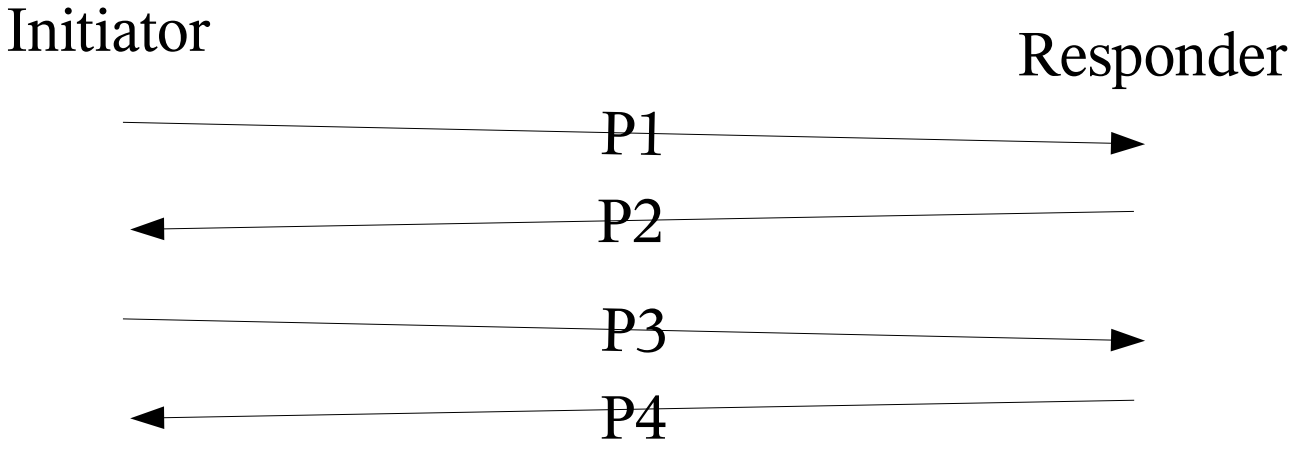
- None
  - A FQDN might be for a service or for a host
  - The FQDN lookup returns a set of potential ULIDs which will be tried by the application until one is working
  - Then the peer will pass its set of locators during the (deferred) context establishment
- Desire to optimize failure during initial contact (by having the multi6 shim try different ones instead of the ULP/application) makes this more complex

Deferred Context Establishment

- Three events occurring at different times
  - Initial contact e.g., some TCP connection to a peer
  - Deciding to setup multi6 context state
    - Based on local policy – port numbers, #packets sent, etc
  - Rehoming the connection after a failure
- Also need to handle failures during the initial contact
  - Base case: punt to the application layer to try different ULID
  - Possible to optimize by having shim do something?

Context establishment exchange

- No state change on receipt of P1
  - DoS protection
- If ICMP error or no response to P1
  - no shim6 support
- Very similar to the HIP exchange



Changes since multi6-l3shim-00.txt (1)

- Using "address" vs. "locator" and "ULID" more consistently and carefully.
- Made it more clear that the ULID is just an IPv6 address.
- In "Renumbering Implications" added text to point out the small probability of there being a problem.
- Extended the assumption about ingress filtering and exit selection.
- Added clarification to MTU implications.

Changes since multi6-l3shim-00.txt (2)

- Clarified what Centrally assigned ULAs can do which regular IPv6 addresses can't do with respect to the DNS.
- Added suggestion from mailing list that one can use different flow label for the communication when ULIDs=locators, and when they are different.
- Listed a few more open issues.

Changes since multi6-functional-dec-00

- None

Open Issues

- Receive side demultiplexing
  - effects packet formats for data packets
- State management
  - how/when is state removed (explicitly? soft state?)
- Packet formats for control protocol
- [APIs for ULP advice]
- [Path maintenance and exploration protocol]
- [...]

Next Steps?

- Pick one approach and work out the details?
- Suggest to pick
  - Use flow label to carry context tag
  - Different flow label after locator change (number picked by receiver)
  - Unilateral removal of shim6 state, plus error message when no state to trigger peer re-establish
  - Control protocol using new IP protocol type
- Alternative would be to explore 8 byte extension header for data packets after failover

Receive side demultiplexing issue

- Receiver needs to be able to correctly rewrite IP address fields before passing to ULP
- Example: ULID A communicates with ULID B and C
  - Later discovers that ULID B has locators B and C, and ULID C has locators B and C i.e., its the same host
  - Locator B fails
  - The peer will receive packets from locator A to locator C
    - Some of which need to be rewritten to ULID B and others which need no rewrite

RSD: prevent receive side confusion

- Each locator is only used with a single ULID
- Means that a host with e.g. 3 prefixes would have 3 ULIDs and 9 locators
  - Each locator is used with only one ULID
- The locator will uniquely identify the ULID at the receiver
- Example: Prefixes P1, P2
  - ULIDs P1|IID1 and P2|IID2
  - Extra locators P2|IID21 and P1|IID12
  - P2|IID21 is remapped to received to ULID P1|IID1

RSD: carry additional info

- Some “context tag” in each packet that needs to be rewritten by receiver
  - The tag exchanged during context establishment
- Where in the packet does it go?
  - Reusing flow label field?
  - A new extension header?
- Former has some complexity due to overloading, but not packet overhead
- Latter implies an extra 8 bytes in the packets after a locator failure

State management

- Coordinated removal of state
  - Ensure that sender knows when receiver might have removed state
  - Sender will know when state needs to be recreated
  - (Plus rule about not rebooting too fast after state loss)
- Unilateral removal plus error message
  - When receiver doesn't find state, send error message
  - Sender recreates state as a result
  - Weaker security; a MiTM which arrives after start can force the setup to be redone

Control Protocol encoding

- Could be IP protocol/nxthdr value
- Could be new ICMP message types
- Could be UDP port number