IPv6 Flow Label for Server Load Balancing - update

draft-carpenter-v6ops-label-balance-02

Brian Carpenter
Sheng Jiang (Speaker)
Willy Tarreau

March 2012
Updated Scenario Diagram

IPv6 Clients in the Internet

Ingress router

DNS-based load splitting

L3/L4 balancer

Possible flow label use

HTTP proxy

HTTP server

HTTP server

HTTP server

TLS proxy

TLS proxy

HTTP server

HTTP server

HTTP server
Use Flow Label to Reduce Work on L3/L4 Load Balancers

- According to RFC 6437, the flow label SHOULD be set to a suitable (uniformly distributed) value at the source.
- Behaviors on a STATEFUL L3/L4 load balancer:
  - if flow label is not present, fall back to current methods (eg: L7 or srcIP)
  - if flow label is present and known, use it to select the proper server. It means, for subsequent packets of flows, the load balancing can depend on 2 tuple \{source address, flow label\} – more efficient for ASICs than 5 tuple today
  - if flow label is present, but new for records, apply the usual load balancing algorithm to select the server and remember the flow label <-> server mapping
- A stateless load balancer can also benefit by using 2 tuple as input of hash algorithm than 5 tuple
Possible Further Improvement: Supporting Session Persistence

- LBs need to maintain session persistence (i.e. always pick the same server) when a transaction includes several transport flows (even different source addresses).
- If applications used the same flow label for all parts of a transaction, LBs could maintain persistence without DPI or session cookies.
  - One flow label per transaction, which may involve multiple transport connections, some of them may from different source addresses.
  - [RFC6437] a flow is not necessarily 1:1 mapped to a transport connection.
- It reduces the work on L7 Load Balancers.
Possible Benefits

- Assuming that 80-90% of users will reach the net without a proxy, large sites will be able to off-load most of their load balancing into ASIC-based LBs or even switches.
  - Ingress router sets flow label if zero
- The remaining 10-20% of sessions will have persistence issues (multiple ports or source addresses) and will follow the normal route via the L7 LBs.
  - Unless we deploy the extended role (same flow label for all parts of a transaction), newly proposed in this document
Questions?

Thanks
Clarification: Who Sets The Label?

- According to RFC 6437, the flow label SHOULD be set to a suitable (uniformly distributed) value at the source.

- Until that becomes general practice, a site using it for server load balancing has two choices when the incoming label is zero:
  - Set the label, per RFC 6437, in an ingress router, thus reducing L3/L4 balancer load except for the first packet.
  - Use the full 5-tuple (as today).
New Security Considerations

- Using a flow label as a transaction handle would require some precautions.
- An unguessable flow label will help in avoiding DDOS attacks on a single server, by making it hard to fool the LB algorithm.
- The LB will store the association between a given flow label value and a given server. This will improve session recovery after a server failure, and also makes it harder for an attacker to target a single server, because this association is not known externally.