IPv6 Flow Label for Server Load Balancing - update

draft-carpenter-v6ops-label-balance-02

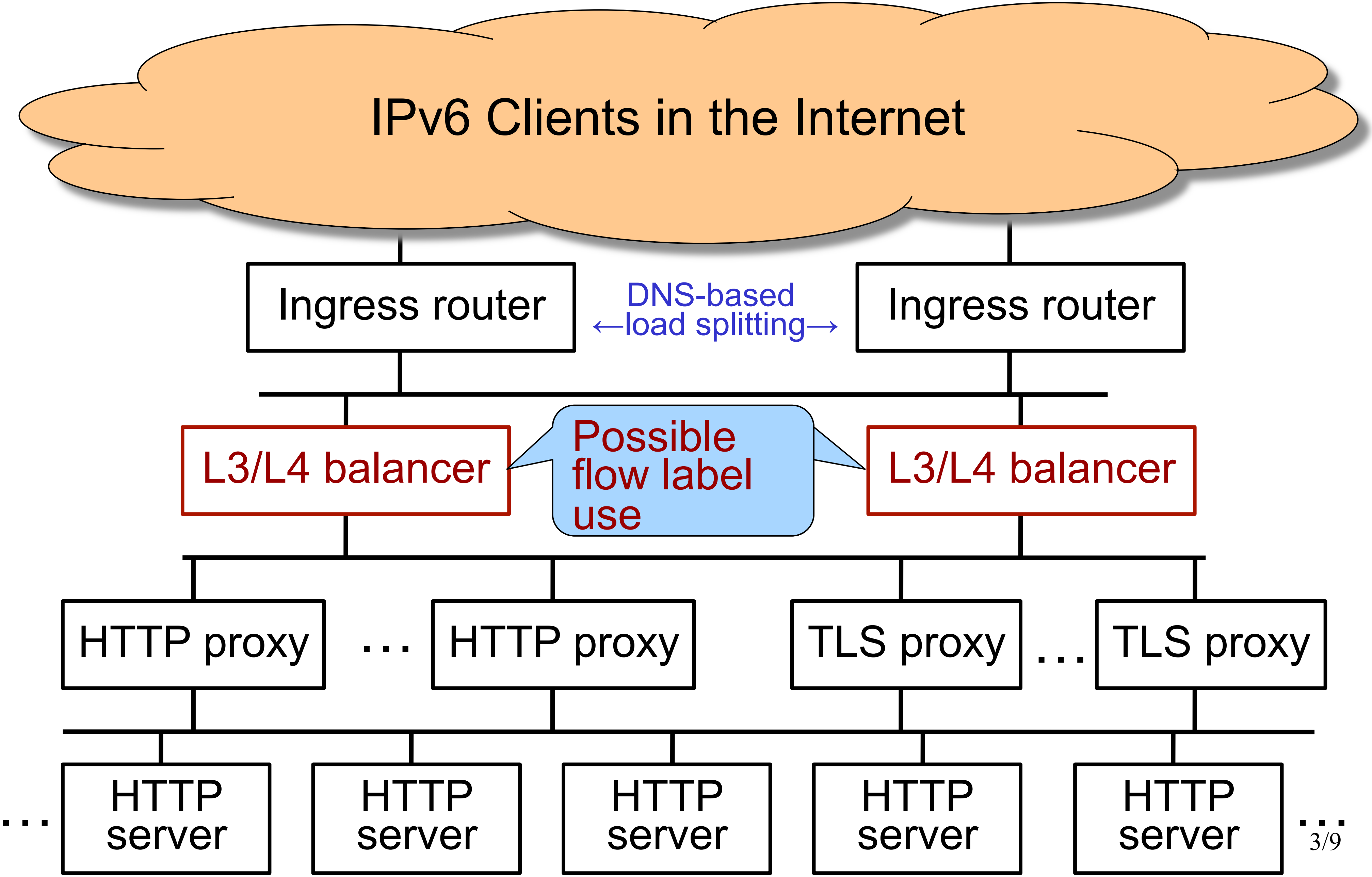
Brian Carpenter
Sheng Jiang (Speaker)
Willy Tarreau

March 2012

Why V6OPS Should Care

- Every serious content provider runs load balancers.
- IPv6 support in load balancers has been a delaying factor for IPv6 deployment.
- The flow label may enhance load balancer efficiency, and even act as an incentive for IPv6 adoption
- No protocol changes – this is implementation and deployment only

Updated Scenario Diagram



IPv6 Clients in the Internet

Ingress router

DNS-based ←load splitting→

Ingress router

L3/L4 balancer

Possible flow label use

L3/L4 balancer

HTTP proxy . . . HTTP proxy

TLS proxy . . . TLS proxy

. . . HTTP server HTTP server HTTP server HTTP server HTTP server . . .

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

- A new flow is directed to a server according to a L7 load balancing algorithm. The flow label doesn’t help there.
- In subsequent packets, the flow label is immediately available regardless of extension headers – more efficient for ASICs.
  - The 3-tuple {source address, destination address, flow label} would be sufficient to identify a *transport* flow, replacing the traditional 5-tuple
  - It can be reduced to 2 tuple {source address, flow label} since destination address is always the same

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).

Use Flow Label to Reduce Work on L7 Load Balancers

- LBs need to maintain session persistence (i.e. always pick the same server) when a transaction includes several transport flows (even different source addresses)
  - Passive-mode FTP picks a new port number.
  - Sessions mix HTTP and HTTPS.
  - Clients behind a web proxy with a dynamic address pool.
- 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

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.

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?

Does the WG want to take on this topic?

Thanks