A Unified Approach to IP Segment Routing

draft-bryant-mpls-unified-ip-sr

Stewart Bryant <stewart.bryant@gmail.com>

Xiaohu Xu <xuxiaohu@huawei.com>

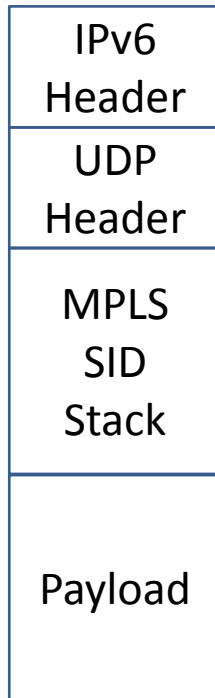Mach Chen <mach.chen@huawei.com>

Adrian Farrel <afarrel@juniper.net>

John Drake <jdrake@juniper.net>

IETF-99, Prague, July 2017
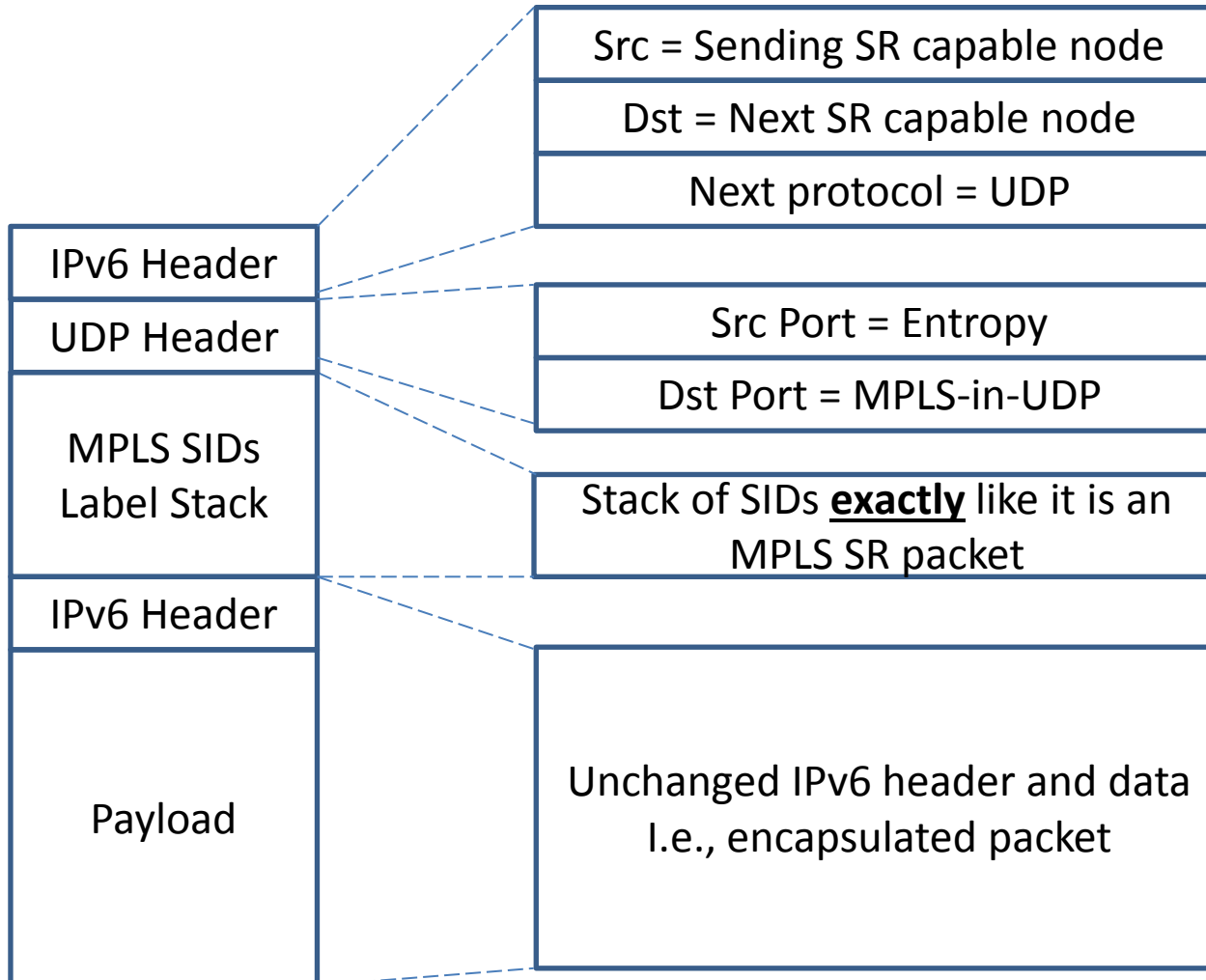
Objectives

1. Tunnel MPLS-SR over an IP network
    - To connect two MPLS-SR networks (e.g., data centers)
2. Produce a unified approach to SR
    - Option of using same MPLS SID stack in all networks
        - Offer same (similar) processing in silicon
    - Support phased deployment and roll-out
3. Co-exist with MPLS-SR and SRv6
    - We appreciate that different solutions are better in different environments and we would "let the market decide"

Overview

- In summary, this is MPLS-over-UDP as RFC 7510

IPv6 Header
UDP Header
MPLS SID Stack
Payload

- Encapsulate a "normal" MPLS SID stack in UDP in IPv6

- Address to next SR-capable node in the SR path

- UDP destination port indicates "MPLS below"

A Little More Detail

IPv6 Header
UDP Header
MPLS SIDs Label Stack
IPv6 Header
Payload

Src = Sending SR capable node
Dst = Next SR capable node
Next protocol = UDP

Src Port = Entropy
Dst Port = MPLS-in-UDP

Stack of SIDs **exactly** like it is an MPLS SR packet

Unchanged IPv6 header and data I.e., encapsulated packet

Advertising SIDs

- Advertisements are just like for MPLS SR
  - IGP or BGP advertises
    - Address of node or link
    - Associated SID
  - All SID types are supported
  - Nothing new is needed

Source Processing

- Build and impose MPLS-SR stack
- Encapsulate in UDP
  - Dst Port = MPLS-in-UDP
  - Src Port = Entropy
- Encapsulate in IP with
  - Source as this node
  - Destination address of first hop in SR stack
    - Requires look-up to match SID to address
    - At source, this lookup can use RIB, etc.
- FIB lookup and send
- (This all looks a lot like RFC 7510)

Transit Non-SR Processing

- Important that this mechanism can traverse nodes that are not SR-capable
  - Also, no special processing by SR nodes to which the packet is not addressed
- It's just an IP packet, so forward it
- ECMP entropy is achieved through the UDP source port value set by source
  - Established technique (RFC 7510)
- TTL decrements as usual

Transit SR Processing

- If the packet is addressed to me
  - Otherwise just forward the packet as normal IP
- Find UDP inside
- Find UDP Dst port is "MPLS-in-UDP"
- Look at top of MPLS SR stack
  - Extract SID and look up "next hop" IP address
  - Pop label stack entry
- Re-encapsulate packet as MPLS-in-UDP-in-IP (just as source did)
  - IP Src = this node
  - IP Dst = next address as found from label lookup
  - UDP Src Port = Entropy (ideally from received packet)
  - UDP Dst Port = MPLS-in-UDP
  - (SID stack is "shorter")
- FIB lookup and send

Use Cases
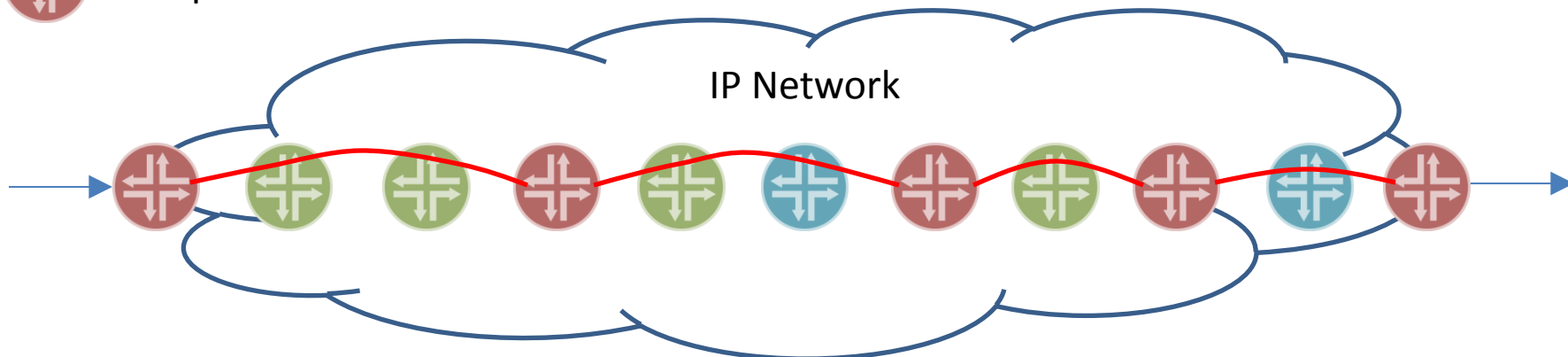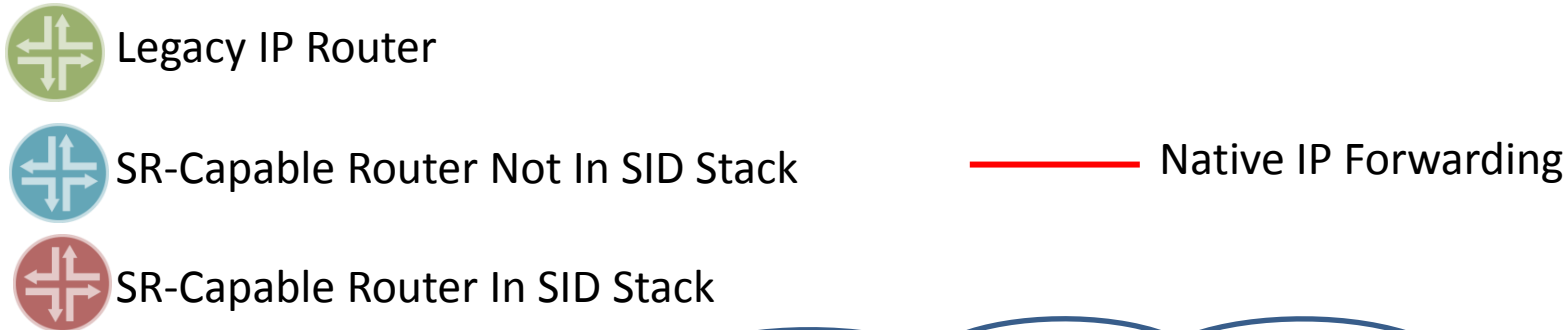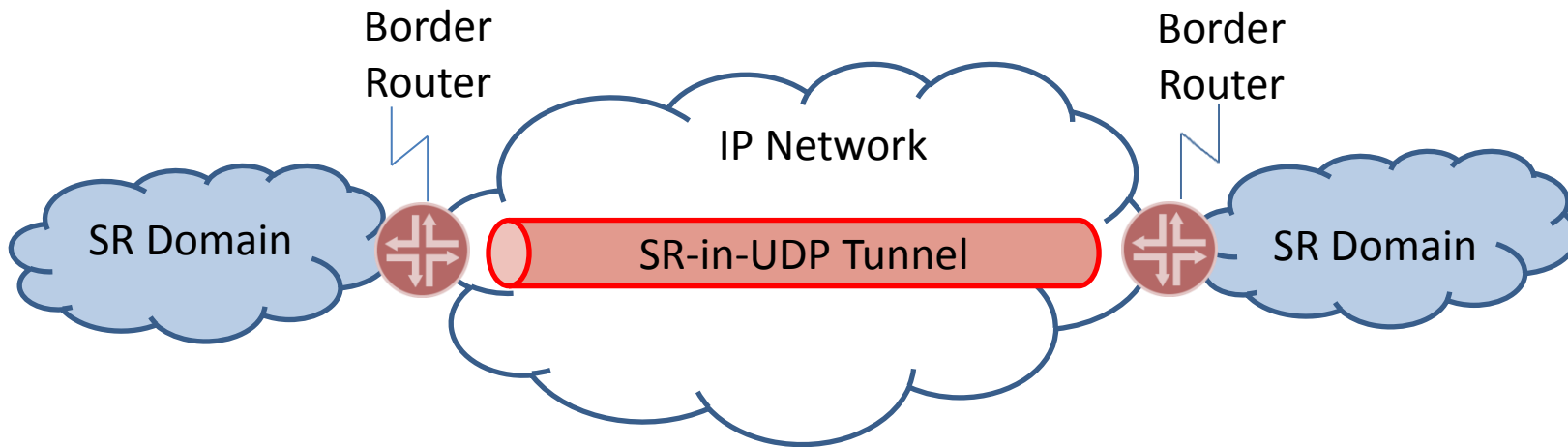
Next Steps

- This is a -00 draft
  - We know some things need polish
- To the authors it seems "obvious"
  - What do other people think?
- The authors think this is in charter for MPLS
  - But polish and discussion is needed before we ask for adoption