draft-ravisingh-mpls-el-for-seamless-mpls-00

Entropy Label for Seamless MPLS

Ravi Singh (ravis@juniper.net)
Yimin Shen (yshen@juniper.net)
John Drake (jdrake@juniper.net)

IETF-86 (Orlando)
speaker: Ravi Singh
Entropy label: Summary (from RFC-6790)

- RFC-6790 defines EL usage for
  - Single-segment LSPs

- Router roles:
  - Ingress LER: label-inserting-router computes and PUSHes the (ELI+EL) before PUSHing the transport-LSP label
    - Based on egress signaling its ability to POP (ELI+EL)
  - (Egress LER)/PHR: removes the (ELI+EL) after POPing transport-LSP label
  - Transit router: utilizes presence of the ELI to hash solely on the EL/label-stack to pick outgoing link of ECMP/LAG

- Use of (ELI+EL) reduces maximum-payload of the LSP by 8 bytes
Terminology:

- Intrinsically-EL capable/capability (ELC): An LSP (segment or e2e) is (or has) ELC: when...
  - ingress LER of (segment or e2e) LSP: has ability as specified by RFC-6790
  - egress/PHR of the (segment or e2e) LSP: has ability as specified by RFC-6790

- Notional ingress/egress LER: Ingress/egress LER (usually a stitching point) for an LSP segment that is respectively PUSHing/POPing the (ELI+EL) on traffic going over an e2e LSP

- Notional LSP segment: portion of the e2e LSP between a consecutive notional ingress and notional egress LER. An e2e LSP might have more than 1 such.
Entropy label & Seamless MPLS:  

- Seamless MPLS:
  - Deploy MPLS in access/aggregation networks
  - By setting up e2e LSPs

- E2e LSPs rely on:
  - LSP stitching
  - LSP hierarchy

- On an e2e LSP:
  - LAG/ECMP might appear anywhere between the ultimate ingress and ultimate egress
  - Routers on path from ultimate ingress to ultimate egress will have varying hash computation capabilities
Entropy label & Seamless MPLS:

- Aim: Get load-balancing benefits of EL wherever possible
- How to get load-balancing benefits…
  - when the e2e LSPs may not be intrinsically ELC?
  - when a transit router’s forwarding ASICs not able to include an EL (past a certain depth in the label stack) for hashing?
  - without unnecessarily reducing the payload capacity of the e2e LSP?
- **This draft**: extends/optimizes EL definition for LSP stitching and LSP hierarchies. Specifies…
  - rules of ELC propagation at stitching points;
  - data-plane guidelines at the stitching point; and
  - the data/mgt-plane guidelines for LSP hierarchies for inserting (ELI+EL) at ingress LER.
LSP stitching: Problems / requirements

- LSP stitching point involves:
  - Incoming LSP segment: Li
  - Outgoing LSP segment: Lo
  - A stitching point router that is connecting Li and Lo

e2e LSP is made of LSP segments L1, L2, L3 and L4.
LSP stitching: Problems / requirements 2/2

- Problem: How to...
  - get load-balancing benefits even though an e2e LSP may not be intrinsically ELC?
    - How to get EL benefits even though the e2e LSP may not support ELC from end to end?
  - not run into data plane issues due to EL insertion?

- Requirements: Modes to be supported:
  - Per-segment ELC
  - ELC for notional segment LSP(s)
  - ELC for e2e LSP
LSP hierarchy: Problems / requirements

- Problems: How to...
  - prevent unnecessary reduction of max-payload of the LSP by EL?
  - prevent possibility of EL being unusable?

---

Entropy Label for Seamless MPLS
IETF-86

8
LSP hierarchy: Problems / requirements

- Requirements:
  - Insert only 1 (ELI+EL) on a data packet
  - Flexibility in choice of LSP tunnel for which EL is inserted
LSP stitching: New abstractions

- Additional router roles:
  - Notional ingress: the segment ingress that is inserting (ELI +EL). Could be different from e2e LSP ingress.
  - Notional egress: the segment egress that is POPing (ELI +EL). Could be different from e2e LSP egress.

- Possibility of multiple notional ingresses and notional egresses on an e2e LSP.
  - 1 notional ingress per notional egress
  - Notional ingress and notional egress alternate

- EL lifetime on a packet:
  - Between a notional ingress and a notional egress
LSP stitching: New abstractions

**SIGNALING:** ELC translation rules at stitching point: translating ELC from Lo to Li

- As segment egress…
  - signals ELC…
    - when this egress segment is intrinsically ELC for Li, or
    - when segment Li is not intrinsically ELC, but segment egress for segment Lo is ELC.
  - MUST NOT signal ELC when…
    - Lo’s segment egress is not signaling ELC, and
    - this router does not have ability to POP (ELI+EL) on Li.

- As segment ingress:
  - Bidir LSPs: When this router is signaling ELC on Li, must also signal ELC for Lo
FORWARDING: Data plane rules to be implemented:

- Handling differing EL dispositions at stitching point:
  - Notional egress/ingress behavior: Differing (Li, Lo) ELC will result in insertion/removal of (ELI+EL) for traffic going into Lo for following (Li, Lo) ELC cases-space:
    - {(no-ELC, ELC), (ELC, no-ELC)}
  - Implicit notional ingress behavior: when (Li, Lo) ELC is (ELC, ELC), this router will insert (ELI+EL) if incoming data packet does not have (ELI+EL)
  - Preventing multiple (ELI+EL) on a packet: insert (ELI+EL) only if incoming packet does not already contain an (ELI+EL)
  - Dealing with role changes due to configuration
    - “Notional ingress -> Not a notional-ingress”, or vice versa
    - “Notional egress -> Not a notional egress”, or vice versa
LSP Hierarchy: New abstractions

Ensuring EL stays usable for load-balancing:
- Management plane triggered data plane changes:
  - Allow disabling insertion of (ELI+EL) on a per LSP basis
  - Allow tweaking of the LSP in the hierarchy that ends up with an (ELI+EL) inserted

Preventing multiple ELs getting imposed on a data packet:
- A router MUST NOT insert an (ELI+EL) on a data packet that already contains an ELI
Next steps

- Discussion on the MPLS WG mailing list
- Moving the draft towards acceptance as a WG document