Running MPLS efficiently in ring networks

Ravi Singh (ravis@juniper.net)

NANOG-65 [Montreal]

Agenda

- Recap: Network-reality, MPLS, rings
- MPLS in rings: Challenges
- Addressing the challenges: Resilient MPLS Rings (RMR)

Recap: Network-reality, MPLS, rings

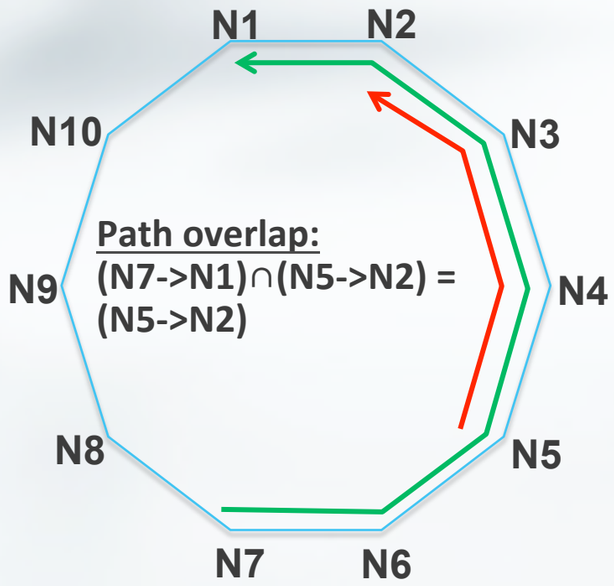
- Network-operators care greatly for:
  - Uptime
  - Effective network utilization
  - Financial efficiency
    - Equipment cost (CapEx): Judicious hardware choice: Devices(routers), Interconnects (links)
    - Operational ease (OpEx savings): provisioning and management

Recap: Network-reality, MPLS, rings (cont'd…)

- MPLS: Enables the foregoing:
  - Uptime: Fast failure protection (Fast reroute)
  - Effective network utilization:
    - Traffic-engineered paths, multi-pathing(LSP-ECMP), entropy-label
    - LSP Bandwidth management: bandwidth accounting, auto-bandwidth
  - Financial efficiency: Allows service-delivery convergence over a shared MPLS network

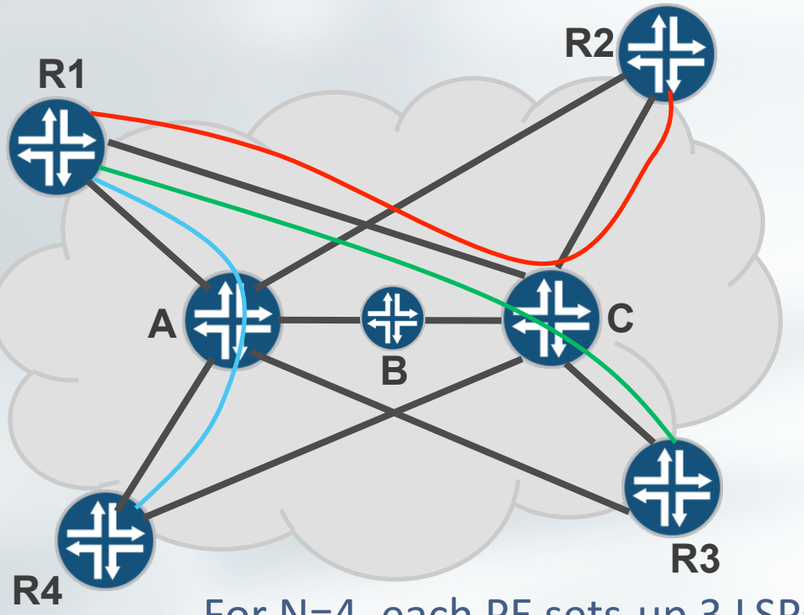
Recap: Network-reality, MPLS, rings (cont'd…)

- Rings' properties cater towards financial efficiency, in choosing a network topology:
  - Adequate # of interconnects to connect N nodes with some failure protection
  - For a given node-pair:
    - 2 paths between the nodes:
    - # of paths reduced to 1 by single failure
    - Second or subsequent failure partitions the network
  - Among node-pairs:
    - Path-overlap
    - Path-overlap leads to bandwidth-preciousness
- **Deployed where the above is acceptable:** *access & aggregation*



MPLS in rings: General challenges

- RSVP-TE signaled LSPs:
  - For complete connectivity
    - Provisioning overhead: $O(N^2)$ LSPs
    - LSP state in network:
      - For $O(N^2)$ LSPs: ok for path-diversity
      - Additional state for bypasses
  - After failure:
    - Inefficient path (over bypass): until ingress LER re-computes path
  - Signaling characteristics:
    - After failure, global repair signaling on a per-LSP basis
    - No exploitation of shared paths to setup MP2P LSPs to a given egress



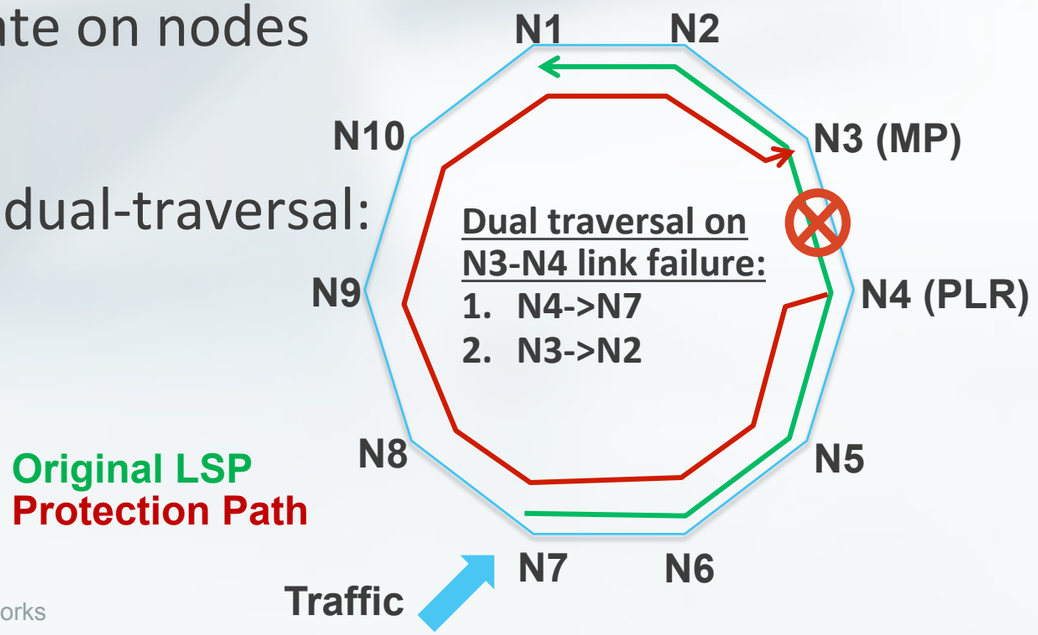
For N=4, each PE sets-up 3 LSPs
Total: 12 LSPs [N*(N-1) LSPs]

MPLS in rings: General challenges (cont'd…)

- LDP signaled LSPs:
  - No support for
    - Traffic-engineered paths with constraints
    - Bandwidth accounting
  - Failure protection:
    - Insufficient failure protection coverage using LFA/RLFA in some topologies
    - Too much path computation churn for alternate paths (RLFA)

MPLS in rings: Ring-specific challenges

- Rings: A peculiar topology for MPLS complete-connectivity
- Ring properties => expose opportunities to make MPLS more effective in rings:
  - Excessive # of LSPs: $O(N^2)$ LSPs excessive:
    - Excessive: At most 2 paths between a node-pair. All LSPs are path-overlapping.
    - $O(N^2)$ LSPs' control-plane & data-plane state on nodes
    - $O(N^2)$ provisioning of LSPs in network
  - Wasted bandwidth: Failure-protection causes dual-traversal:
    - From point-of-local-repair (PLR) to ingress
    - From merge-point (MP) to egress



MPLS in rings: Ring-specific challenges (cont'd…)

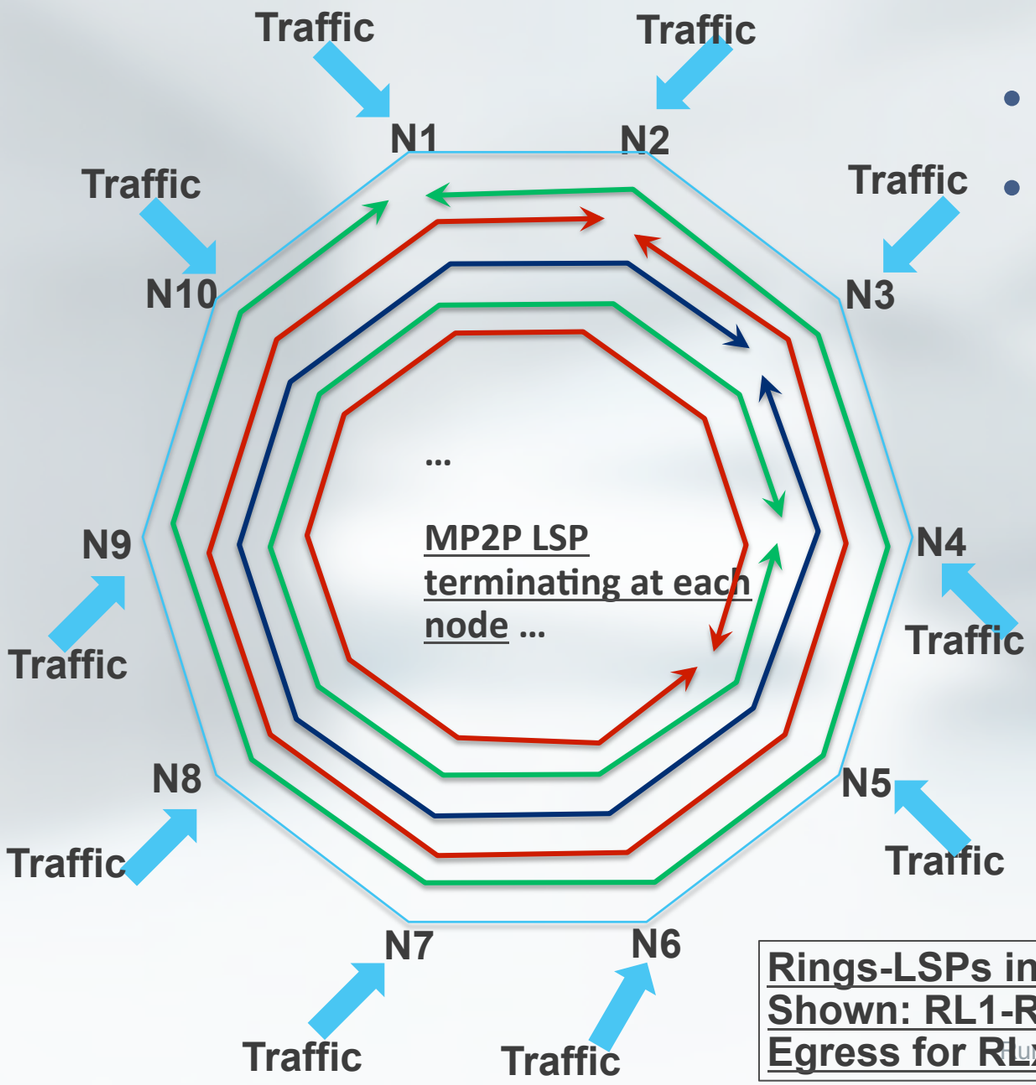
- Ring properties => Optimizations to enable efficiency in running MPLS

- General-purpose MPLS challenging to use:
  - Too generic in trying to signal LSPs and their protection paths
  - Complete topology coverage: Leads to excessive LSP state than strictly necessary
  - Does not:
    - Give optimal # of LSPs given significant path overlap
    - Utilize local-protection effectively
    - Utilize ring bandwidth effectively

Resilient MPLS Rings (RMR): Benefits

- Concept similar to SDH BLSR (Bidirectional Line Switched Ring)
- All LSPs benefit from:
  - Automatic ring discovery
  - Self-configuring rings
- RSVP-TE signaled LSPs:
  - Easier provisioning
  - Fewer LSPs: lesser control-plane state, less onerous processing, smaller FIB
- LDP signaled LSPs
  - More effective failure protection

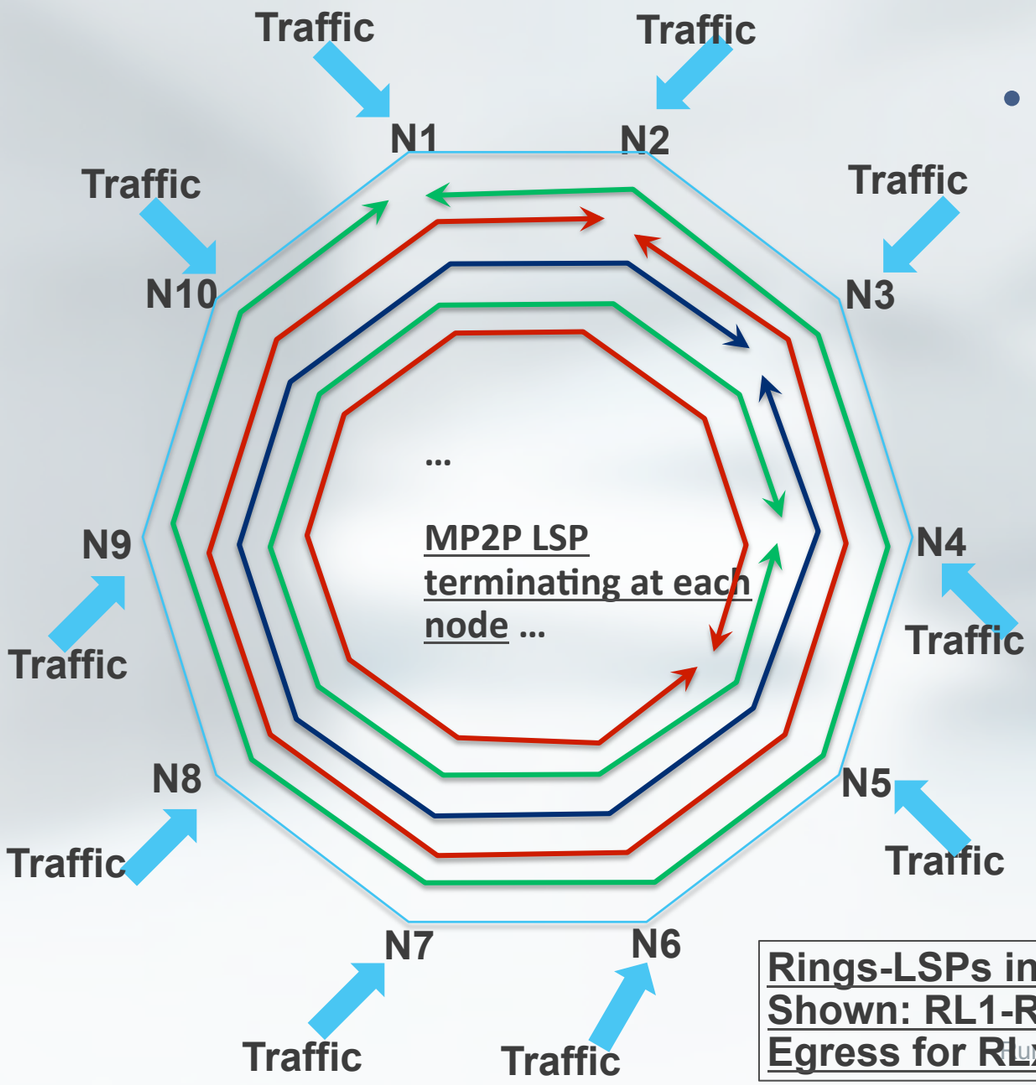
RMR: LSP characteristics



- A ring of N nodes has N ring-LSPs
- Each ring-LSP is:
  - MP2P (has a single egress)
  - Made of two unidirectional sub-LSPs
    - Clockwise (CW), and
    - Anti-clockwise (AC)
  - Either sub-LSP (not both) can carry given flow
  - Any ring-node can send traffic on a ring-LSP
  - Can have varying LSP-bandwidth along its various hops

Rings-LSPs in this ring: RL1 to RL10
Shown: RL1-RL5
Egress for RLx is node Nx

RMR: LSP characteristics (cont'd…)



- Protection:
  - Involves switching traffic from affected-direction sub-LSP to the other-direction-sub-LSP
  - Does not use bypasses/detours/etc.

Rings-LSPs in this ring: RL1 to RL10
Shown: RL1-RL5
Egress for RLx is node Nx

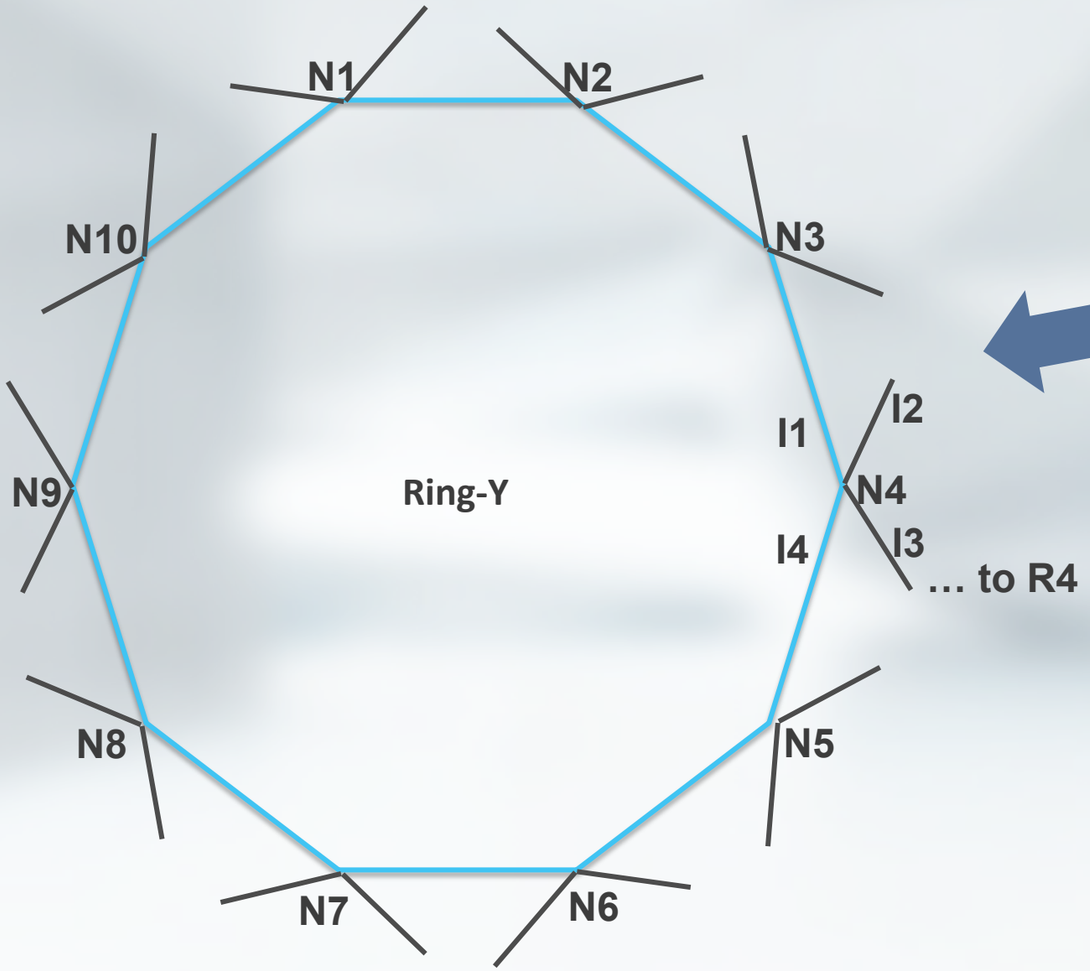
RMR: Easier provisioning

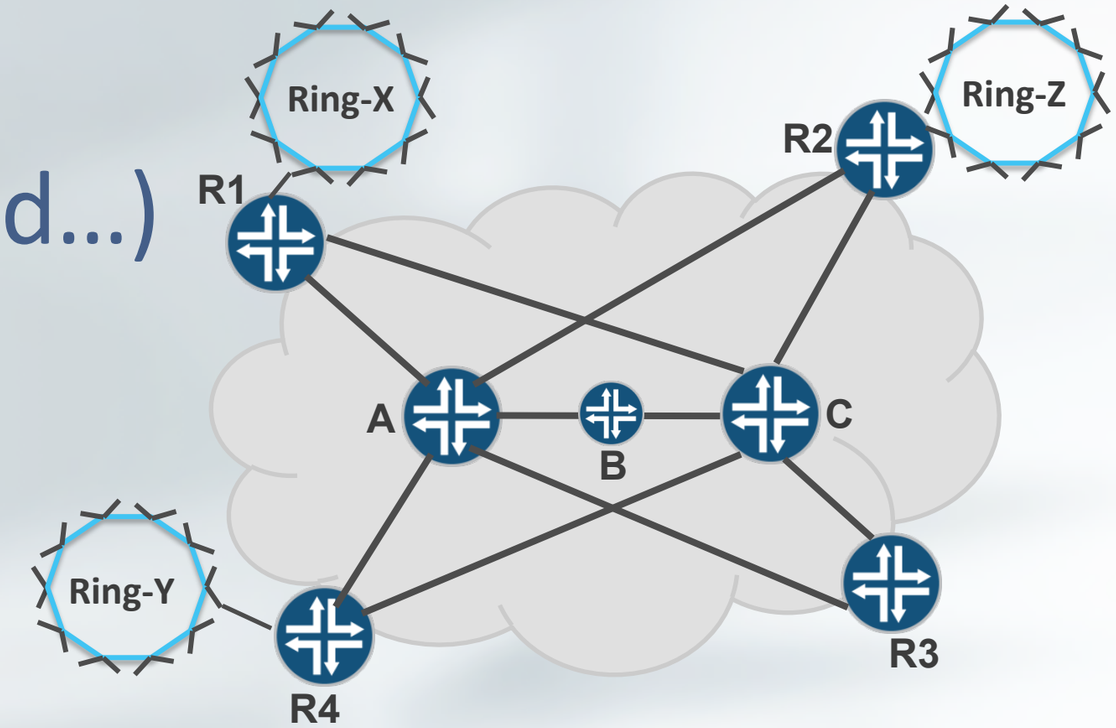
- Ring ID:
  - Configure ringIDs on relevant interfaces of subset of nodes
  - Do not have to configure ringID on every node in ring
  - Promiscuous nodes: nodes not configured with ringID acquire ringID of neighbor node(s)

- Configuration focus on ring and not an LSP
  - Configure rings and not LSPs: so configuration overhead at most O(N)
  - Ring LSPs automatically set themselves up

- Works over ring-links configured as unnumbered links

RMR: Easier provisioning (cont'd…)





- Configuring interfaces on N4:
  - set interface I1 ring-link ring Y
  - set interface I4 ring-link
- I1 configured for specific ring
- I4 configured for promiscuous mode
- Configure IGP, RSVP/LDP to run on I1/I4
- No per-LSP configuration

RMR: Ring auto-discovery

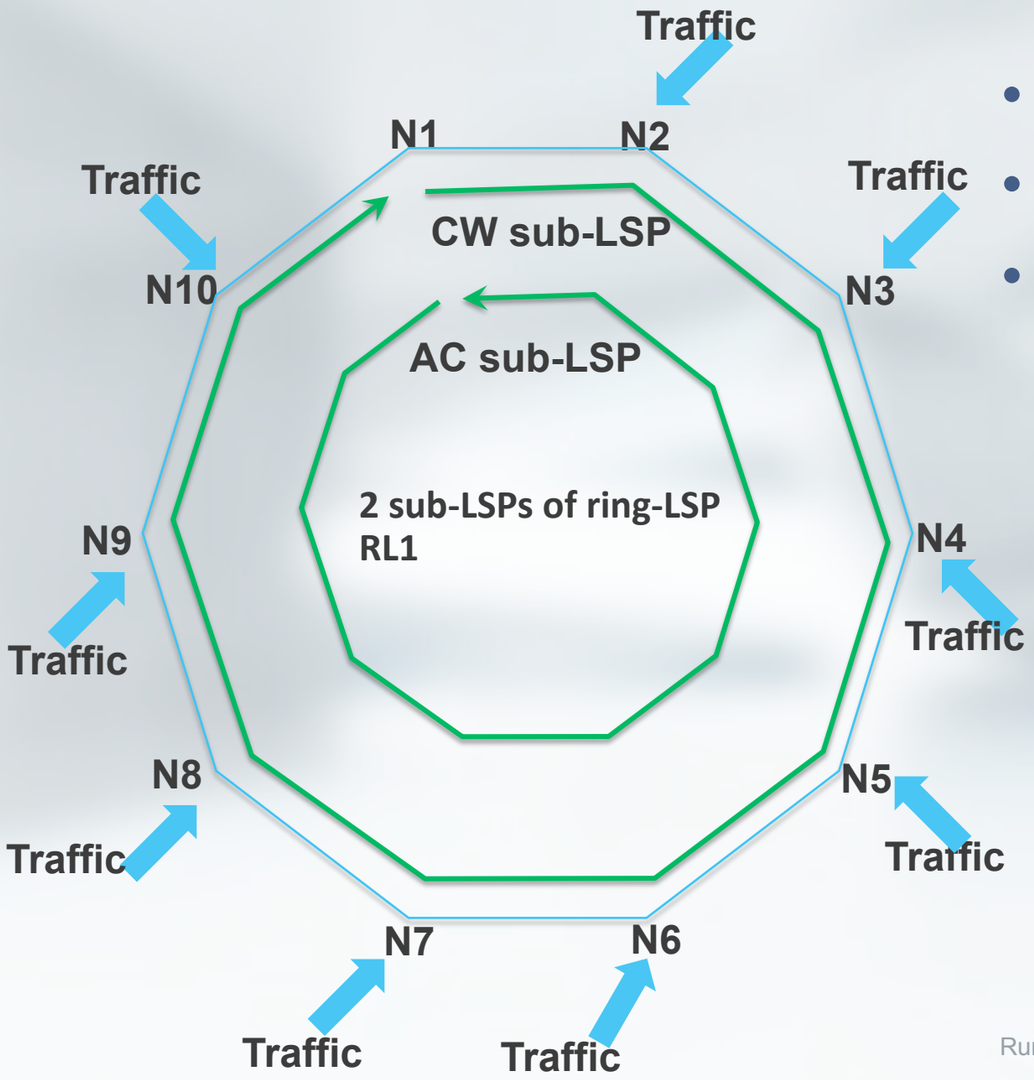
- Ring information propagated through an IGP to
  - Allow ring-nodes in same ring to discover each other
  - Allow for (re)discovery of a maximal ring topology when node/links added/deleted
  - Signal to all nodes in the ring, the ordered list of all nodes participating in the ring in both directions

- Pre-cursor to signaling ring-LSPs

RMR: Ring auto-discovery (cont'd…)

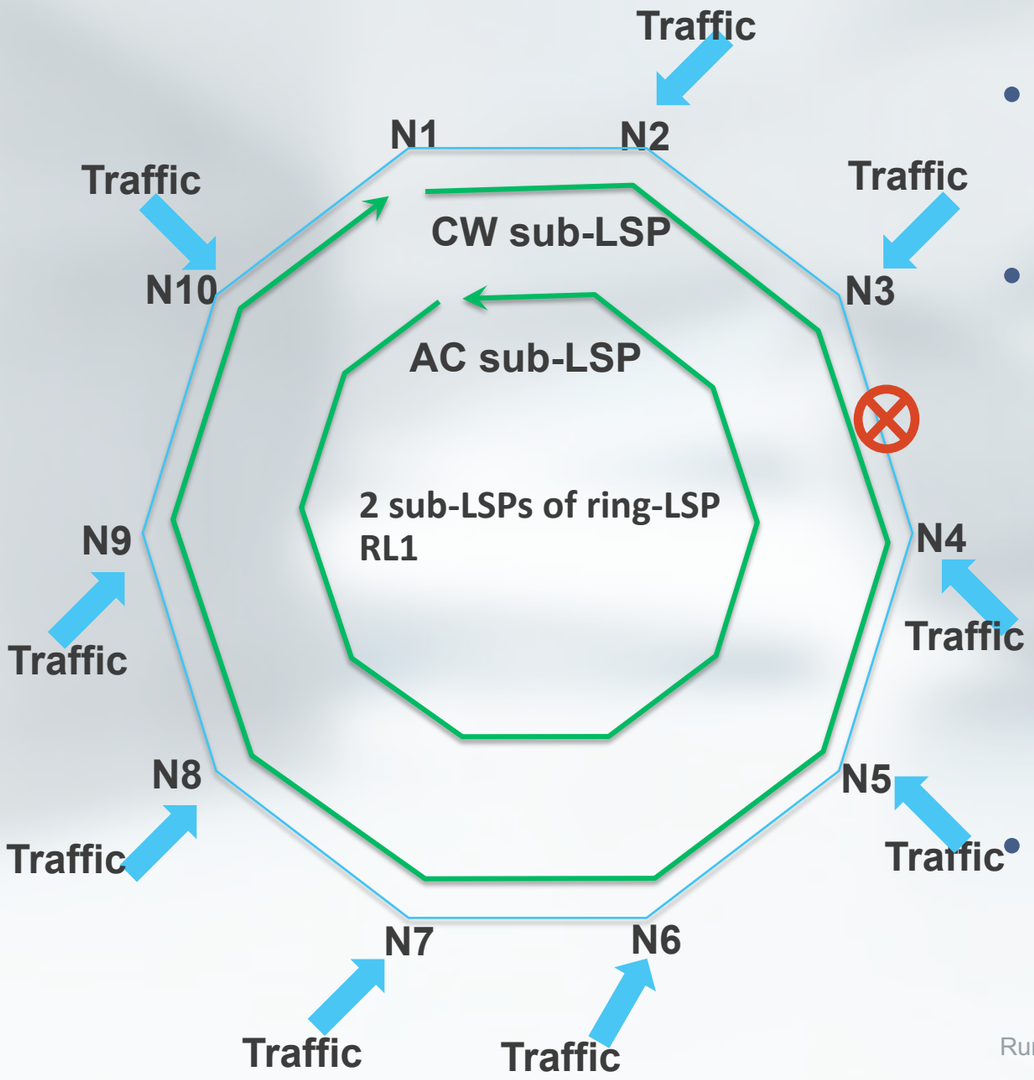
- Three phases:
  - Announcement phase: Configured ring-nodes advertise ringID. Promiscuous nodes pass same on.
  - Mastership phase:
    - Ring-master elected to administer ring-discovery
    - Ring-master decides when the ring-discovery is deemed completed
  - Ring identification:
    - Initiated by ring-master on election
    - Concluded by ring-master on it acquiring AC & CW neighbors for the ringID

RMR: Ring in steady-state



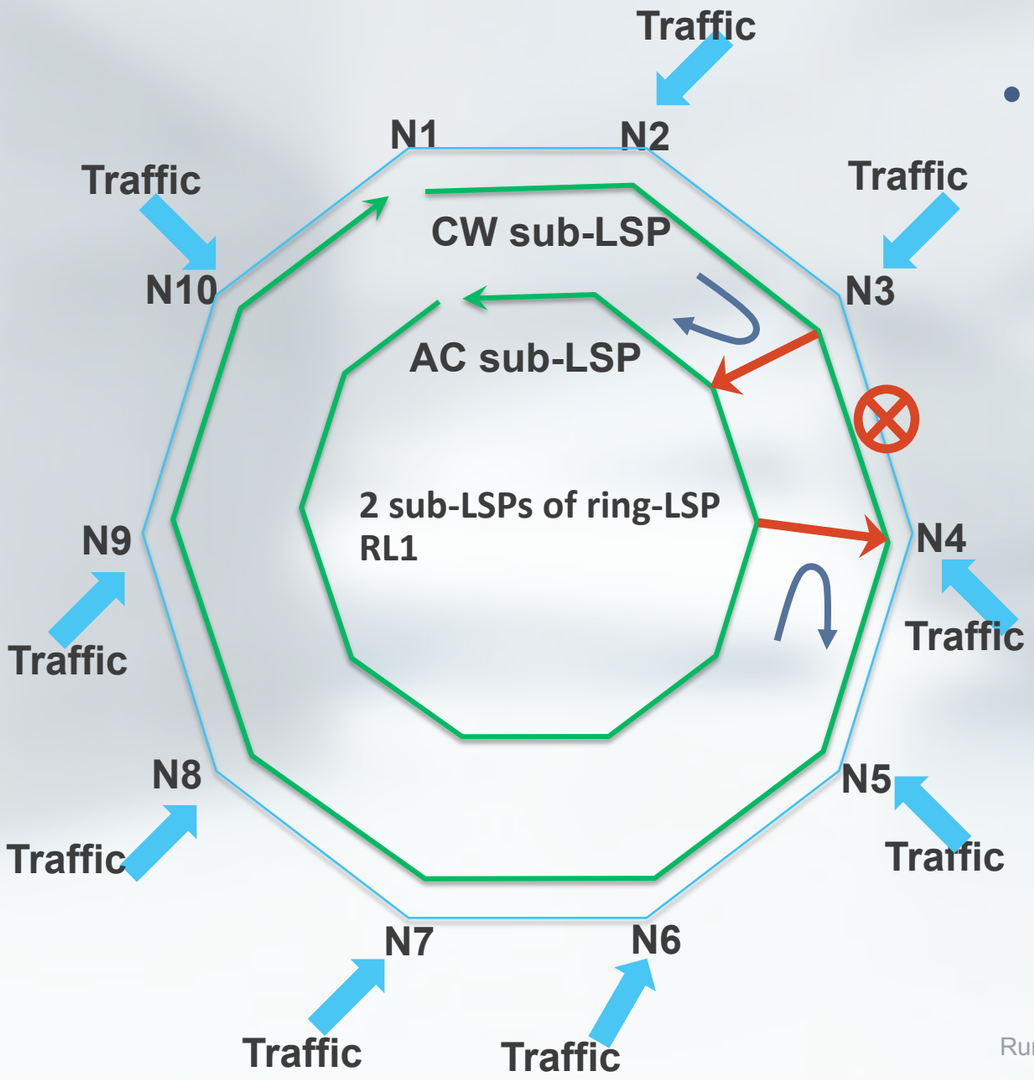
- Using RL1 for illustration (Same applies to all LSPs)
- RL1 has N1 as egress node
- 2 sub-LSPs: CW & AC
  - Both terminate at node N1
  - Any node N2 to N10 may
    - Send traffic for a given flow to N1 on either CW or AC sub-LSP
    - Utilize both sub-LSPs at the same time

RMR: Ring during failure



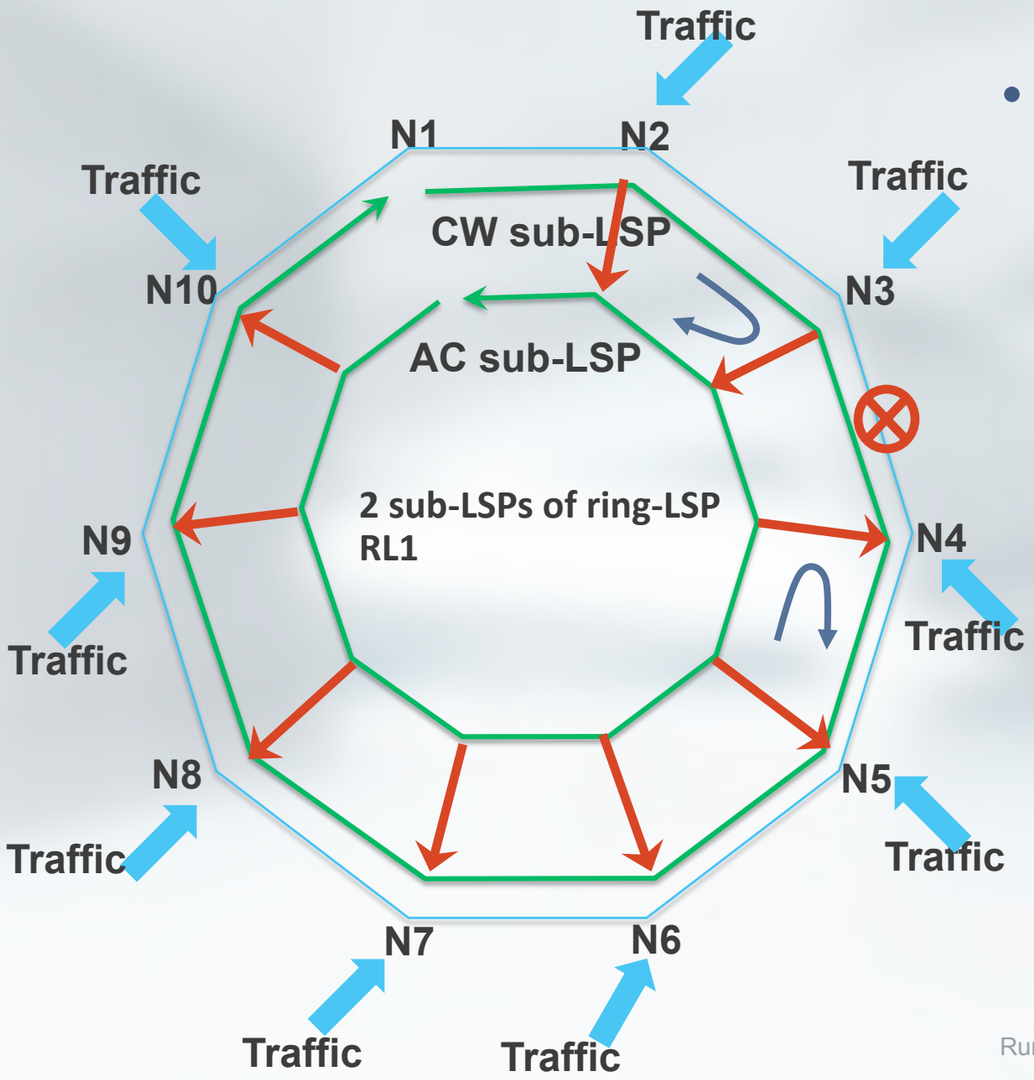
- Using RL1 for illustration. Same applies to all LSPs
- LSP failure protection
  - Local repair by failing-over to the counter-rotating counterpart sub-LSP
  - Global repair: simply propagating local-repair upstream (on other direction LSP) up to ring-LSP egress
- Say link between N3 & N4 fails:

RMR: Ring during failure (cont'd…)



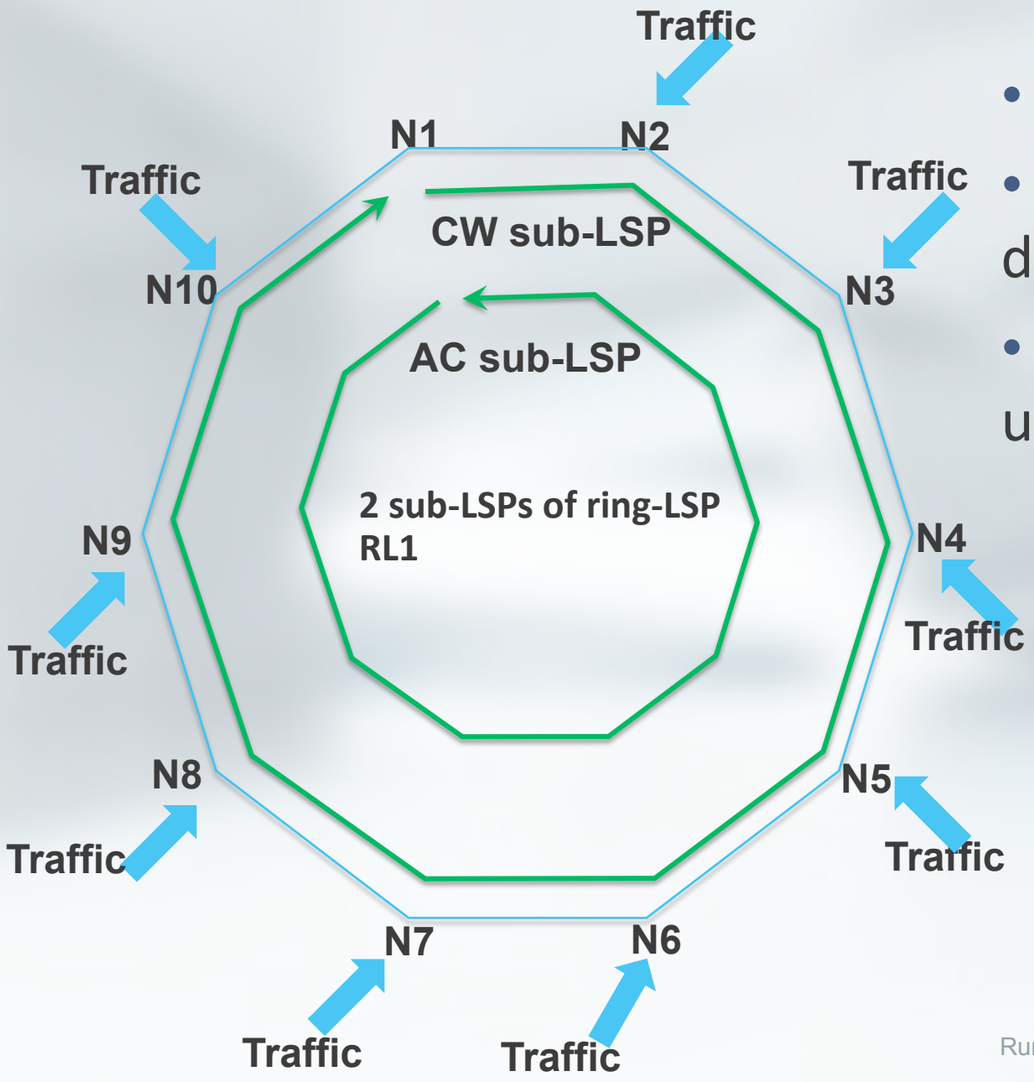
- Post-”failure of link N3-N4”: Local repair
  - At N4: AC-sub-LSP switches to CW-sub-LSP
  - At N3: CW-sub-LSP switches to AC-sub-LSP

RMR: Ring during failure (cont'd…)



- Post-”failure of link N3-N4": Global repair
  - No re-signaling of RL1 (i.e. of AC-sub-LSP and CW-sub-LSP)
  - Error propagation on each sub-LSP causes cascaded traffic-switch to the other sub-LSP
    - Switching to AC sub-LSP in order: N3, N2
    - Switching to CW-sub-LSP in order: N4, N5, N6, N7, N8, N9, N10

RMR: Bandwidth management



- Sub-LSPs of a ring-LSP are MP2P
- Path overlap dictates that various hops will have different amount of bandwidth reservations
- Differing per-hop bandwidth reservations and utilizing both sub-LSPs facilitate:
  - Fewer LSP preemptions (compared to regular MPLS) when various ingresses increase bandwidth requirements

RMR: Comparison to regular MPLS:

- RSVP-signaled:
  - Provisioning overhead:
    - Much reduced
    - Configure ringIDs and not individual LSPs
  - LSPs are self-provisioned by ring-discovery
  - # of LSPs: O(N) as compared to $O(N^2)$
  - MakeBeforeBreak (MBB):
    - Only to adjust per-hop bandwidth of LSP
    - No change in labels of LSP during MBB
  - During global-repair: No path re-computation at ingress
- LDP-signaled:
  - Complete ring coverage for protection
  - No micro-loops by design
  - Prior to failure can use both paths

RMR: Features enabled & protocol extensions

- Features enabled:
  - Auto-hierarchy
  - Automatic LAG traversal
  - Flexible load-balancing at ingress
  - More complex aspects:
    - Multiple rings
    - Rings of rings
- Protocol extensions:
  - OSPF/ISIS:
    - Ring discovery extensions
  - RSVP-TE:
    - Ring-LSP 5-tuple definition
    - Per-hop-varying LSP-bandwidth constructs
  - LDP
    - Ring-LSP TLV definition

Summary

- RMR makes MPLS easier to deploy in rings by:
  - Reducing configuration overhead: ring auto-discovery
  - Reducing LSP state in the network thus allowing scale-down of hardware specs (CPU/memory)
  - Enabling more effective bandwidth utilization in the ring
  - Enabling more effective failure protection
    - By reducing link dual-traversal
    - By giving complete coverage for the entire ring

References

- Resilient MPLS rings: https://tools.ietf.org/html/draft-kompella-mpls-rmr-01

Thanks

V1.10