Segment Routing benefits and learnings

Steve Dyck
Principal Consulting Engineer
Nokia

Why Segment Routing ?

1

Simplify

Consolidate protocols

Reduce network state info

Simplify multi-domain networking

Simplify network operations

Increase network scale

2

Control

Network programmability

Centralize policy control

Traffic engineering

- bandwidth
- latency
- bi-directionality
- path diversity
- load balancing, …

Flow steering / service chaining

3

Automation

Path visualization

Real time monitoring / rebalancing

Bandwidth on demand / proactive rebalancing

Path re-optimization

Segment Routing standardized by IETF

Driven by SPRING Working Group

Source Packet Routing In NetworkinG

- First draft 2013
- Over 97 docs and counting
  - 38 published RFCs
  - 32 drafts
  - 27 working group docs

- https://www.segment-routing.net/ietf

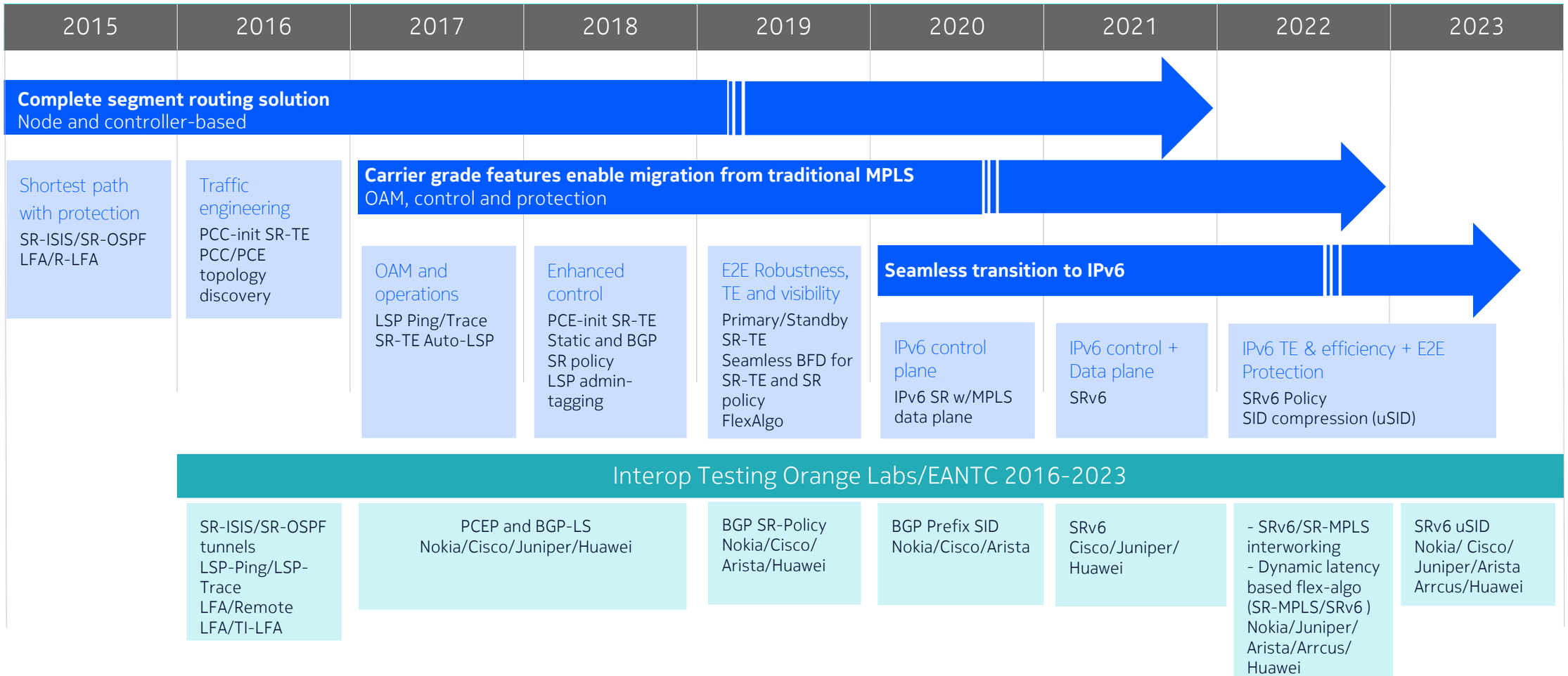
Goal: SDN for wide area networks

Scope includes

- Architecture
- SR-MPLS, SRv6
- Use-Cases & Requirements
- Deployment & Interop
- Fast Reroute
- Network Resource Partitioning (entropy, slicing, …)
- OAM, Path Trace, Performance Measurement
- BGP, IS-IS, OSPF
- PCEP
- Replication
- Yang models

NOKIA

Segment routing roadmap

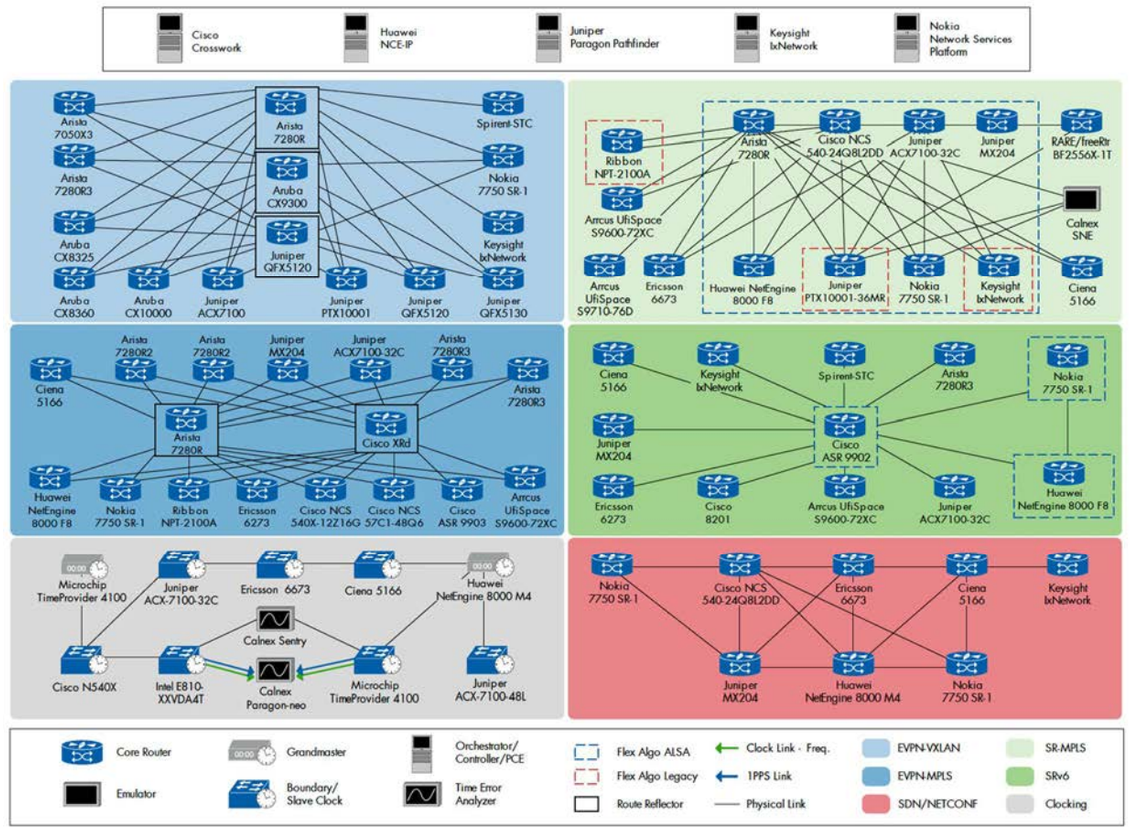


SRv6 Multivendor interop testing

EANTC 2023 – SRv6 Base-SID/Micro-SID

Nokia, Arista, Arrcus, Cisco (Micro-SID only), Huawei, Juniper, Keysight and Spirent test cases:

- EVPN-VPWS MH AA over SRv6 → Cisco, Huawei, Nokia, Keysight and Spirent
- EVPN-VPWS SH over SRv6 → Cisco, Huawei, Nokia and Keysight
- EVPN-MPLS SH over SRv6 → Cisco, Huawei, Nokia , Keysight and Spirent
- EVPN-L3 (RT5) over SRv6 → Arista, Cisco, Huawei, Nokia and Spirent
- GRT IPv4/IPv6 over Base SRv6 → All vendors participated
- L3VPN IPv4/IPv6 over Base SRv6 → All vendors participated
- SRv6 Ping/Trace → All vendors participated
- TI-LFA with SRv6 → Arcus, Cisco, Huawei and Nokia
- SRv6 and MPLS Service Interworking
  - IPVPN MPLS to EVPN L3 SRv6 → Cisco, Nokia and Arista as IGW
- SRv6 locator aware summarization → Cisco, Huawei, Juniper and Nokia as L1/L2 routers
- Flexible algorithms with SRv6 → Huawei, Juniper, Nokia and Spirent
  - Optimized on delay metric
  - Optimized on TE metric + Include-All AG



Segment routing toolkit

Protection/OAM	LFA/R-LFA/ TI-LFA	Primary-secondary SR-TE/SR-policy	LSP ping/trace seamless BFD
Programmatic control	PCEP	BGP/BGP-LS	NetConf
Traffic engineering	SR-TE	SR-Policy	Flex-Algo
Control plane	ISIS, OSPF, BGP or static, IPv4 or IPv6 Node SID, Adj SID, Adj SET, Anycast SID, BGP Prefix SID		
Data plane	MPLS Label push/pop/swap/forward	IPv6 IP forwarding with SRH (SR Hdr)	

Advanced feature set

Simple, fast forwarding

Segment IDs

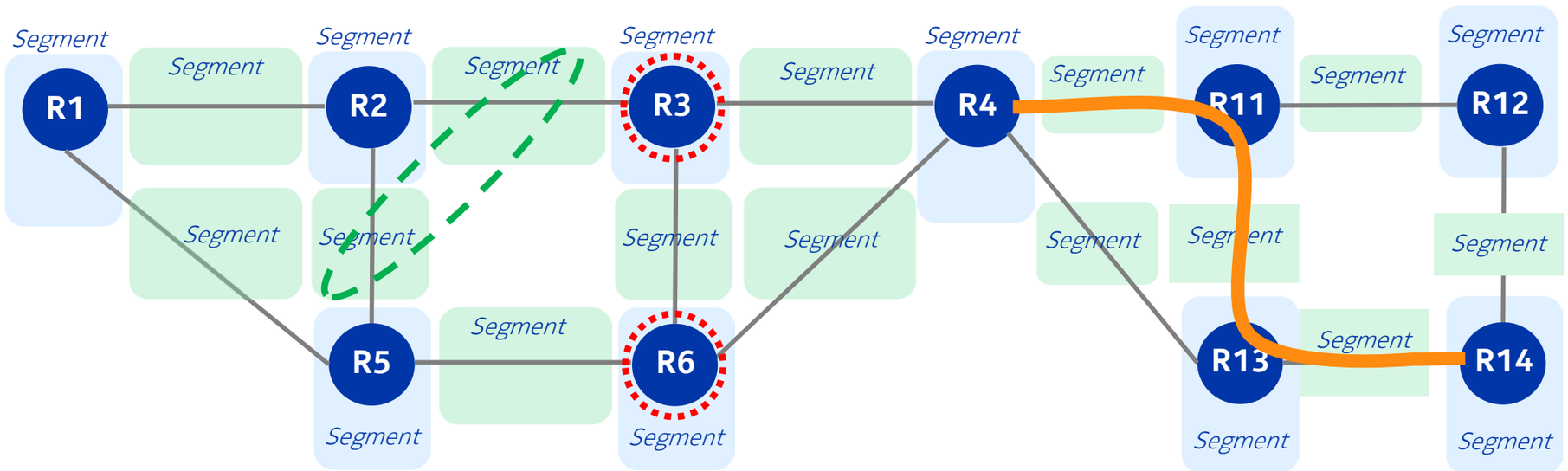
Node SID

Adjacency SID

Adjacency Set

Anycast SID

Binding SID



Segment Routing tunnel can follow the shortest path ECMP-aware* route through a network or follow a specific set of instructions (segments).

The instructions can be programmed using a centralized controller, or programmed in a distributed manner with each SR-capable router autonomously computing the Constrained Shorted Path First (CSPF)

*If ECMP enabled, the source router will create multiple tunnels to endpoint and spread traffic across those tunnels. Entropy label can also be inserted to allow hashing along the path.

Max label stack depth is configurable. As per RFC 3032 (MPLS Label Stack Encoding), each label is 4-bytes.

Flexible Algorithms

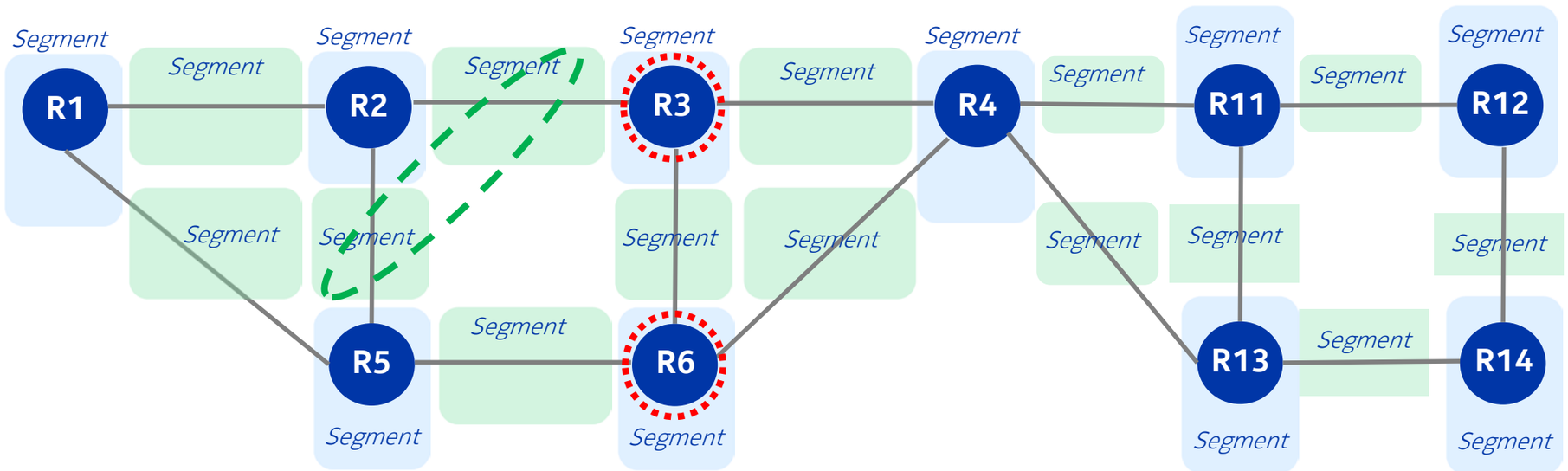
Node SID

Adjacency SID

Adjacency Set

Anycast SID

Binding SID



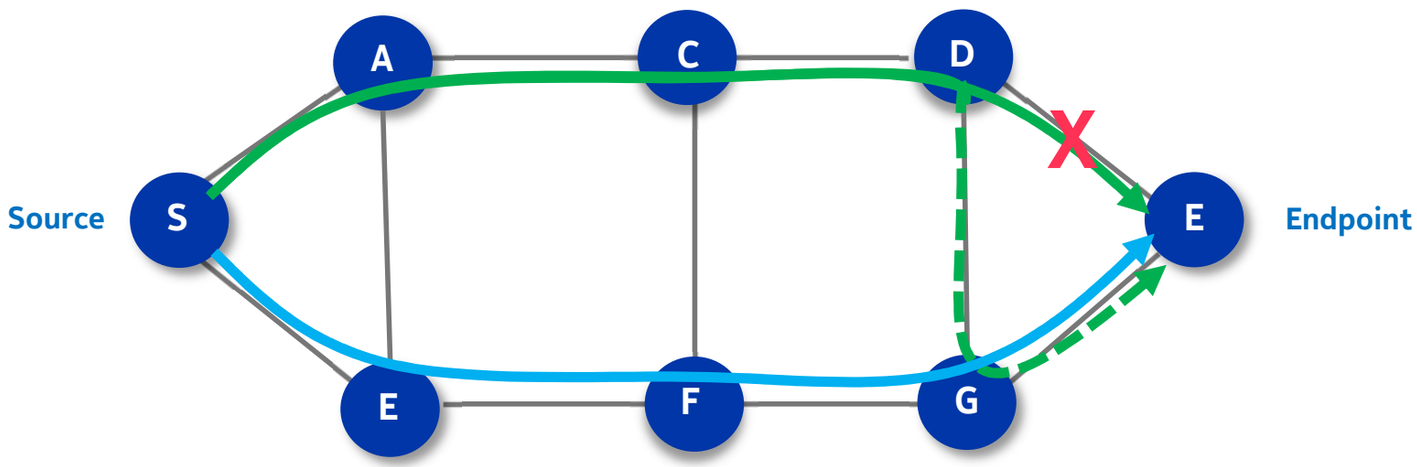
Flex Algo: Create Different topology views of the physical network

1) Select which metric to use for SFP computation: IGP metric, TE metric or Delay Metric
2) Select which links to consider for SFP computation:
   - Exclude <list of colors> and/or
   - Include ANY <list of colors> and/or
   - Include ALL <list of colors> and/or
   - Exclude SRLG <list of colors>

Resiliency for SR-TE tunnels

Fast Reroute

- Use LFA, Remote LFA or Topology Independent LFA
- Use pre-computed alternate next hop as local point of repair
- Once source node learns of failure (BGP-LS) recalculate post-convergence primary path (2-5 sec)



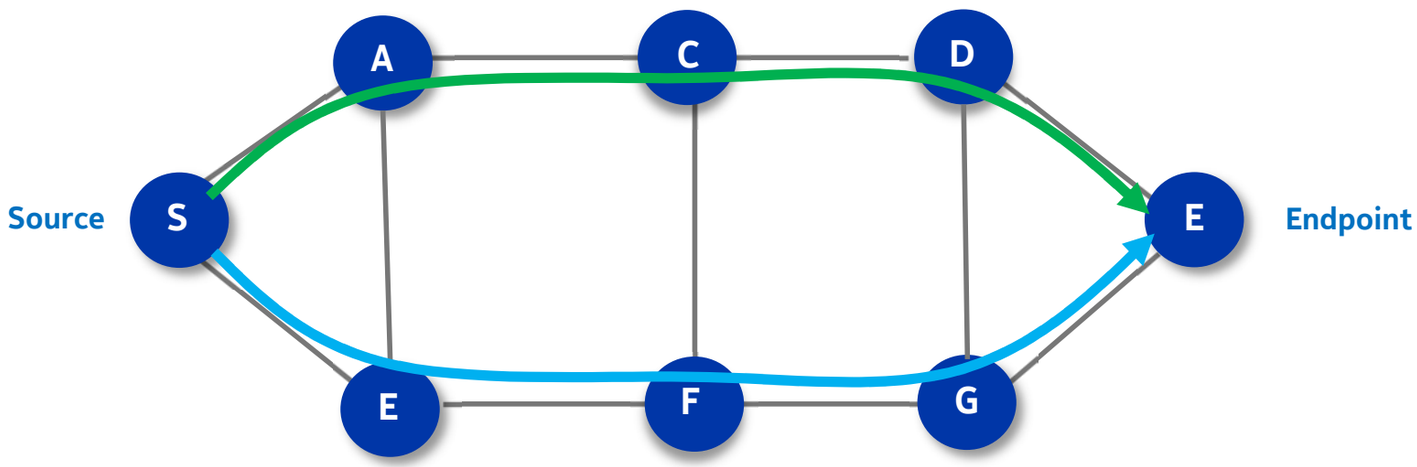
Resiliency for SR-TE tunnels

Fast Reroute

- Use LFA, Remote LFA or Topology Independent LFA
- Use pre-computed alternate next hop as local point of repair
- Once source node learns of failure (BGP-LS) recalculate post-convergence primary path

Optional: Primary / Secondary disjoint paths

- SR-TE with pre-computed disjoint primary AND standby paths
- Seamless BFD on both paths for end-to-end continuity test to trigger rapid switchover
- BFD granularity to 10 msecs



Path Computation Element

Centralized control and visibility

Real-time network visualization

Improve SLA adherence by engineering paths end-to-end with routing decisions based on

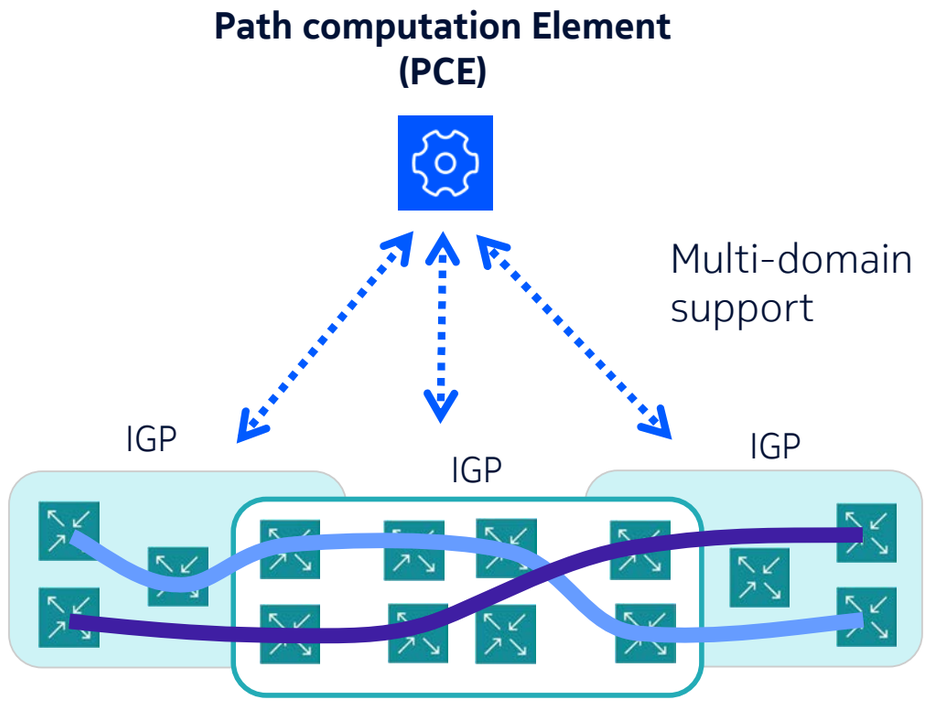
- Actual utilization
- Bandwidth (bandwidth management)
- Latency requirements

Automate path [re-]routing

- to distribute traffic for optimal network capacity usage
- to trigger path re-optimization
- to prevent issues from maintenance actions

Network Resiliency (primary/secondary)

Offline simulation



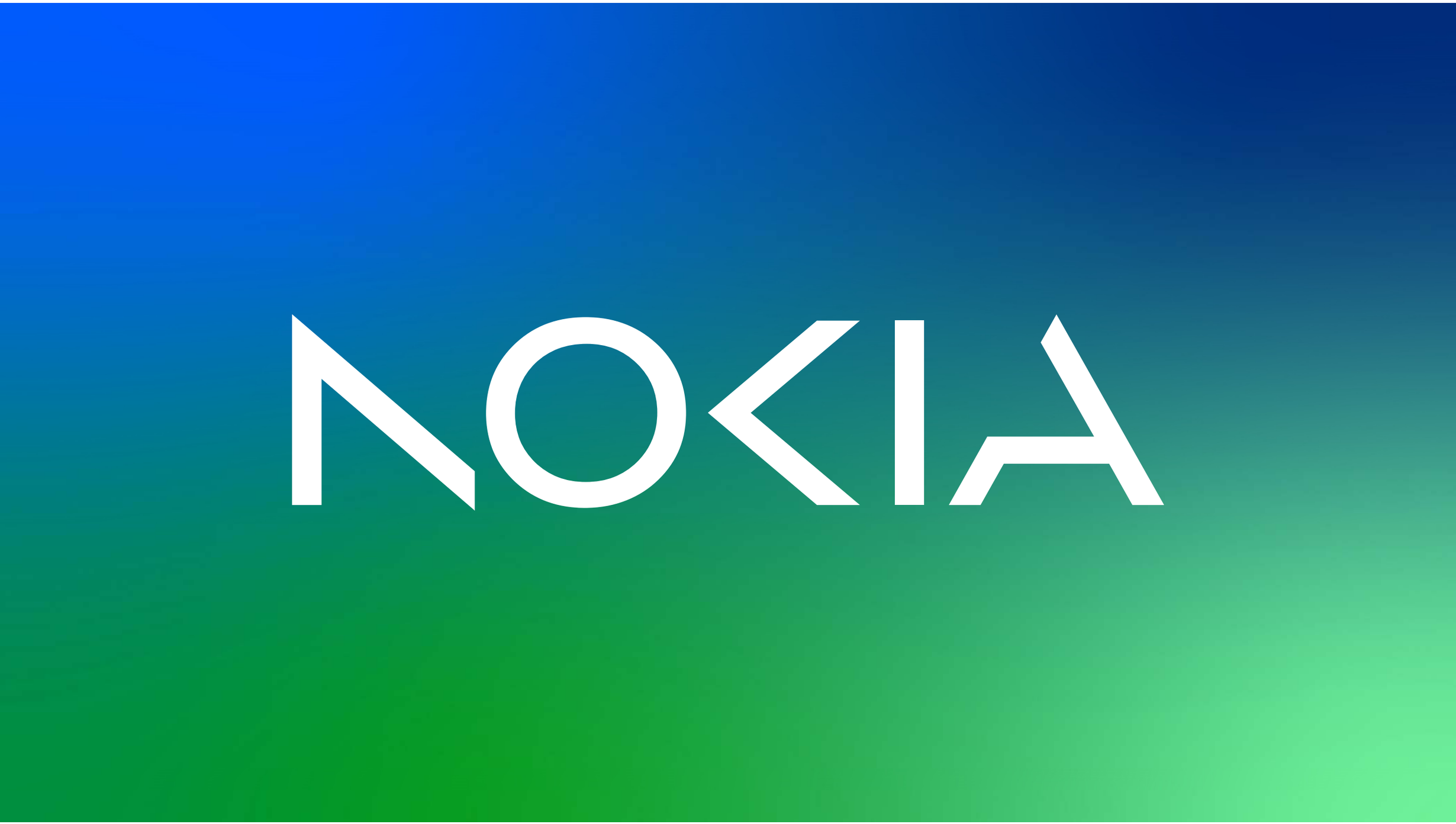
Optimize end-to-end paths to avoid congestion, latency and SLA impact

SR-MPLS and SRv6 Deployment Considerations

	SR-MPLS	SRv6
IP Infrastructure	• IPv4 or IPv6	• IPv6 only
IGP (ISIS/OSPF underlay)	• New ISIS/OSPF extensions	• New ISIS/OSPFv3 extensions
BGP (service overlay)	• Same as traditional MPLS services	• new SRv6 services SID TLV
IP Address Planning	• No impact on existing dual-stack IPv4/IPv6	• Requires one locator prefix per node, per flex-algo
Data Path Bandwidth Efficiency	• Good	• Requires SID compression scheme (IETF standardization in progress)
Multicast NG-MVPN	• New BGP or PCEP based Tree-SID	• New BGP or PCEP based Tree-SID
Scope of Deployment	• Wide: single and multi-domain with seamless MPLS and E2E SR-TE	• Limited: initially introduced to implement VPN overlay for tenants and VNF/CNF applications • E2E SRv6 gated on support in low-end devices and telco cloud apps • Service Gateway to interwork with rest of MPLS network
Ease of Migration	• Immediate benefits: TI-LFA and latency topology using Flex-Algo • Additional capabilities enabled by SR-TE / PCE • Mature state of interoperability (EANTC)	• Requires new address planning • New echo-system for automation and troubleshooting • Standardization (IETF) and interoperability (EANTC) in progress
Network and Application Programmability	• Limited: fixed label size • Applications: TE and service chaining	• Flexible: ability to encode variable size application data in EH or SRH. • Applications: VNF/CNF hosting in coupled overlay/underlay, service function chaining, load-balancer, in-situ OAM

Contact

steve.dyck@nokia.com


NOKIA