Segment Routing IPv6 (SRv6) on Nokia routing platforms

Application note

Contents

Introduction

Segment routing is a powerful technology for overlaying services on a transport network and protecting and engineering network traffic, at high scale and with simplified operations. Segment routing traffic can follow the shortest path, Equal-Cost Multi-Path (ECMP)-aware route through a network, or it can follow a specific set of instructions or segments, providing the ability to steer a packet on a specific path according to a source route.

Segment routing can be deployed with a Multiprotocol Label Switching (MPLS) data plane (SR-MPLS) or an IPv6 data plane (SRv6). SRv6 is emerging in edge telco cloud data center use cases as well as new IPv6 backbone use cases. It provides a powerful IPv6 network programmability framework that takes advantage of the large IPv6 address space.

Some network operators may choose to deploy SR-MPLS based on, for example, their familiarity with MPLS technology, installed equipment base and operations environment. Other operators may gravitate towards SRv6 to take advantage of its additional network programming capabilities.

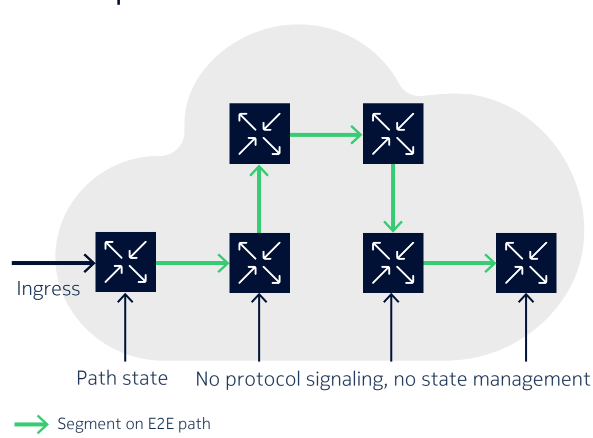
Nokia IP routers enable operators to choose either path with confidence by supporting full-featured implementations of SR-MPLS and SRv6, along with a comprehensive SR-MPLS-to-SRv6 interworking capability. These implementations and capabilities have been proven in public interoperability tests and in real-world production environments.

An accompanying application note, Segment routing on Nokia routing platforms, provides an overview of the operation and capabilities of segment routing and of SR-MPLS support on Nokia routing platforms. It describes the suite of tools and applications provided by the routers to help network operators take full advantage of segment routing capabilities. This application note focuses mainly on the capabilities and benefits of SRv6 on Nokia routing platforms.

SRv6 data plane

Segment routing steers packets by encoding Segment Identifiers (SIDs) in their headers. SIDs contain the complete packet-processing instructions for each intermediate and destination router. This greatly reduces the need for a control plane to instantiate and maintain path state in the network, as illustrated by Figure 1. It also simplifies network operations and reduces resource requirements.

Figure 1. Segment routing end-to-end path and state



Segment routing leverages the IPv6 data plane using a routing header called the Segment Routing Header (SRH). SRv6 SIDs are encoded as IPv6 addresses. A segment list can be formed from a single segment or an ordered list of segments within the SRH.

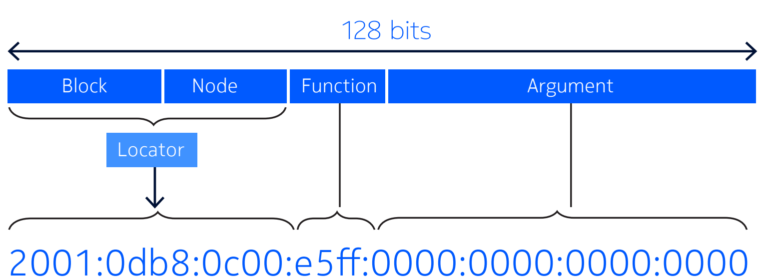
In shortest path routing, the source node originates an IPv6 packet with a single segment encoded in the destination address of the IPv6 header. There is no explicit requirement to include an SRH, except for a Remote LFA (RLFA) or Topology-Independent LFA (TI-LFA) backup path. A source-routed path contains an ordered list of segments encoded as a list of 128-bit IPv6 addresses that represent all the segments in the segment list in the SRH.

SRv6 Segment Identifier

The programmatic nature of SRv6 is supported by the structure of the SRv6 SID, which is divided into a Locator, a Function and an Argument, as shown in Figure 2. Essentially, the Locator is a routable prefix of the destination router that advertised the SID. It consists of an SRv6 subnet block and a node-specific identifier. The Function field encodes a routing, service or network service function context in which the packet must be processed at the router that terminates the SID. The Argument field can bring added information for processing the packet in its specific routing or service context.

The lengths of the Locator, Function and Argument fields are operator-defined but the total length must be less than or equal to 128 bits. A feature of the SRv6 SID is that it is still a conventional IPv6 address and can be used to forward an SRv6 packet using the well-established longest prefix match table lookup against the route of the locator prefix according to Classless Inter-Domain Routing (CIDR) rules.

Figure 2. SRv6 SID structure



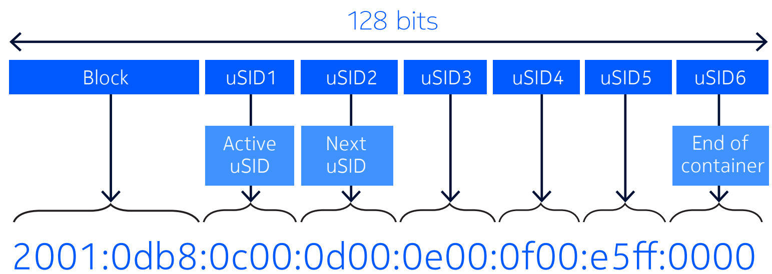
SRv6 SID compression and micro-segment identifier (uSID) instruction

Certain use cases such as traffic engineering require possibly lengthy lists of SIDs in the packet header. The SRv6 SID structure described above is sometimes referred to as a classic SRv6 SID. While acceptable for shortest path applications, the classic SID structure becomes a resource-impacting overhead when the segment list is long. This has a negative impact on bandwidth efficiency and packet processing requirements.

Nokia has been active in standards bodies such as the Internet Engineering Task Force (IETF) and in product development to solve this issue. The SRv6 ‘micro-segment identifier’ (or uSID) instruction is a compressed representation of a segment, and is supported as an extension of the SRv6 network programming model. uSIDs can be used in conjunction with classic SRv6 SIDs within the same segment list. This enables encoding of additional service or network function metadata in a subset of the SIDs within a mostly compressed segment list.

The uSID compression mechanism exploits the fact that a sequence of consecutive SIDs in a segment list usually shares a common Locator Block. The segment list can thus be significantly compressed by avoiding the repetition of the Locator Block and zero trailing bits within each individual SID. A uSID container is a 128-bit SRv6 SID that can be encoded in the destination address of an IPv6 header or at any position in the segment list of an SRH. Figure 3 shows an example of the uSID container structure.

Figure 3. SRv6 uSID container structure



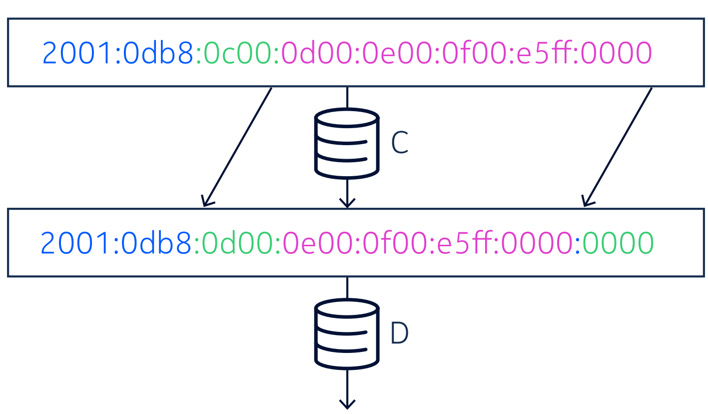
A 32-bit uSID-Block (Locator Block) supports the encoding of up to six 16-bit uSIDs. In the example shown in Figure 3, the container encodes five 16-bit uSIDs plus an end of container marker using a special uSID value of ‘0000.

uSIDs can be global or local. A global uSID is associated with a node ID and identifies a route towards that node. A local uSID identifies an instruction or operation and can be used to identify a local adjacency, a service or a network service function instance. The binding of an instruction to a local uSID is significant within a single node only, although the same uSID value could be reused for other instruction bindings across an SRv6 network domain.

As the packet progresses along the source-routed path, a node terminating the leftmost uSID shifts the container content (except the Block) to the left to expose the next uSID so that the downstream nodes can look it up.

In the example in Figure 4, node C performs a longest prefix match lookup on the container prefix in its forwarding information base (FIB) and hits the route of its local locator 2001:db8:0c00::/48. It then shifts the uSIDs to the left and exposes the next uSID, which is the node ID “0d00” of node D. Router C inserts the end of container marker on the trailer of the container and performs a second-longest prefix match lookup to forward the packet to node D.

Figure 4. SRv6 uSID container shift-and-forward operation



SRv6 toolkit on Nokia routing platforms

Services

The beneficial characteristics of SRv6, including scalability, resilience and operational simplicity, are inherited by overlay services. Nokia routing platforms support a comprehensive set of services with classic (uncompressed) SRv6 and micro-segment SRv6 (uSID).

Supported services and routing contexts include:

- Base router BGP routes
- BGP Virtual Private Routed Network (VPRN) and Layer 3 BGP Ethernet Virtual Private Network Interface-less (EVPN-IFL)
- BGP EVPN Layer 2 Virtual Private LAN Service (EVPN-VPLS) with multi-homing (active/standby or active/active) support
- BGP EVPN Layer 2 Virtual Private Wire Service (EVPN-VPWS) with multi-homing (active/standby or active/active) support

In addition, Nokia routing platforms can use point-to-multipoint SRv6 policy tunnels to efficiently replicate multicast service traffic.

Traffic engineering

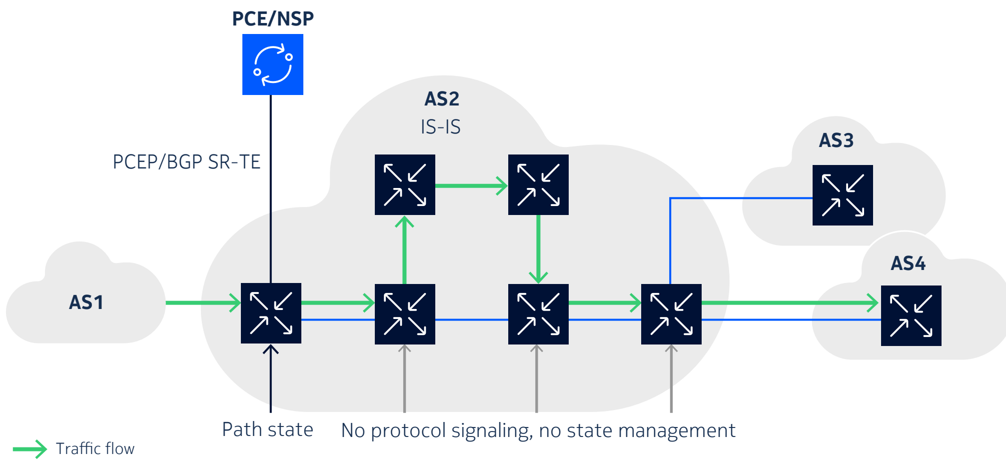
A powerful, yet operationally straightforward, form of traffic engineering can be achieved on SRv6 paths using IGP Flexible Algorithm (Flex-Algo). Flex-Algo can establish an overlay IGP topology to apply constraints to engineer traffic flows, for example, to meet latency targets or to steer service packets through subsets of network links using admin group constraints.

Segment Routing with Traffic Engineering (SR-TE) allows more granular and deterministically defined paths (or policies) to be established using a list of SRv6 SIDs that define a source routed path at ingress. The use of uSID maximizes the value of this application without a significant increase in packet overhead.

SR-TE using SRv6 segment lists enables more advanced applications such as load balancing flows (ECMP or weighted ECMP) across multiple disjointed source-routed paths or steering flows to target service chaining function instances.

Ingress traffic-engineered policies can be defined statically and advertised in a distributed fashion by BGP. With a centralized control plane, a controller known as a Path Computation Element (PCE) is responsible for computing the path, which is then communicated to the node using the Path Computation Element Protocol (PCEP) or BGP using the SR-TE policy address family, as shown in Figure 5.

Figure 5. Path Computation Element and segment routing



The Nokia Network Services Platform (NSP) provides a solution for path control and optimization. PCE functionality is part of the NSP suite and natively integrates with all the other capabilities that the NSP provides for the operational life cycle, including configuration, provisioning and assurance of the network and services. It is possible to operate a network with a distributed control plane and a centralized control plane for continuity if communication with the PCE is lost temporarily.

The NSP controls and optimizes traffic in near-real time with high performance with efficient and scalable path computation and optimization algorithms and a scalable south-bound protocol (IGP, BGP, PCEP) implementation based on the Nokia Service Router Operating System (SR OS). Read more about the Nokia NSP in its role as a PCE in the Path control and optimization in segment-routed networks application note.

Protection

Fast reroute techniques using Loop-Free Alternates (LFAs) provide hardware-based link, node and Shared Risk Link Group (SRLG) fast protection for all SRv6 tunnels using node SIDs and adjacency SIDs (shortest path tunnels and SRv6 policies). RLFA and TI-LFA backup paths take advantage of source routing support in SRv6 to provide fast protection coverage in most of the commonly deployed network topologies.

In addition, Nokia routing platforms support fast, hardware-based, end-to-end protection using Seamless Bidirectional Forwarding Detection (SBFD) to detect data plane failures.

Operations, Administration and Maintenance

SRv6 uses native IP and therefore the classic tools such as Ping and Traceroute provide most of the required troubleshooting and connectivity verification capabilities. SR OS will respond to ping and traceroute targeted at any locally instantiated endpoint.

Nokia routing platforms support SRv6 Operations, Administration and Maintenance (OAM) for the network underlay components (Locator, SID, SRv6 policy) and the service overlay (prefixes resolved to SRv6 tunnel).

SBFD provides an efficient and simplified mechanism with rapid provisioning and improved control and flexibility for path monitoring. It supports fast forwarding path failure detection and uniform failover for SRv6 policy tunnels.

Multi-instance service interworking gateway capability

Nokia supports the industry's most flexible segment routing gateway solution, which provides the ability to interwork and translate among different data and control planes, including MPLS, SR-MPLS, SRv6, VXLAN and MPLS over UDP.

One example of gateway usage is in a selective introduction of SRv6 into a MPLS-based network. MPLS/SRv6 gateways in the base router or in a virtual routing and forwarding (VRF) instance allow a seamless extension of services and connectivity to the rest of the network. This simplifies and eases the introduction of SRv6 and preserves investments already made in the network.

This capability de-risks the decision to choose SR-MPLS or SRv6 technology. A network operator could choose a pragmatic deployment of SR-MPLS now and make a smooth evolution to SRv6 later. Alternatively, different domains within the same network to cater to a specific operational environment or to take advantage of a specific data plane capability. For example, an operator could use SR-MPLS in a WAN deployment and SRv6 in a data center environment. The SRv6 use cases section provides more details on the deployment use cases of the gateway capability.

SRv6 use cases

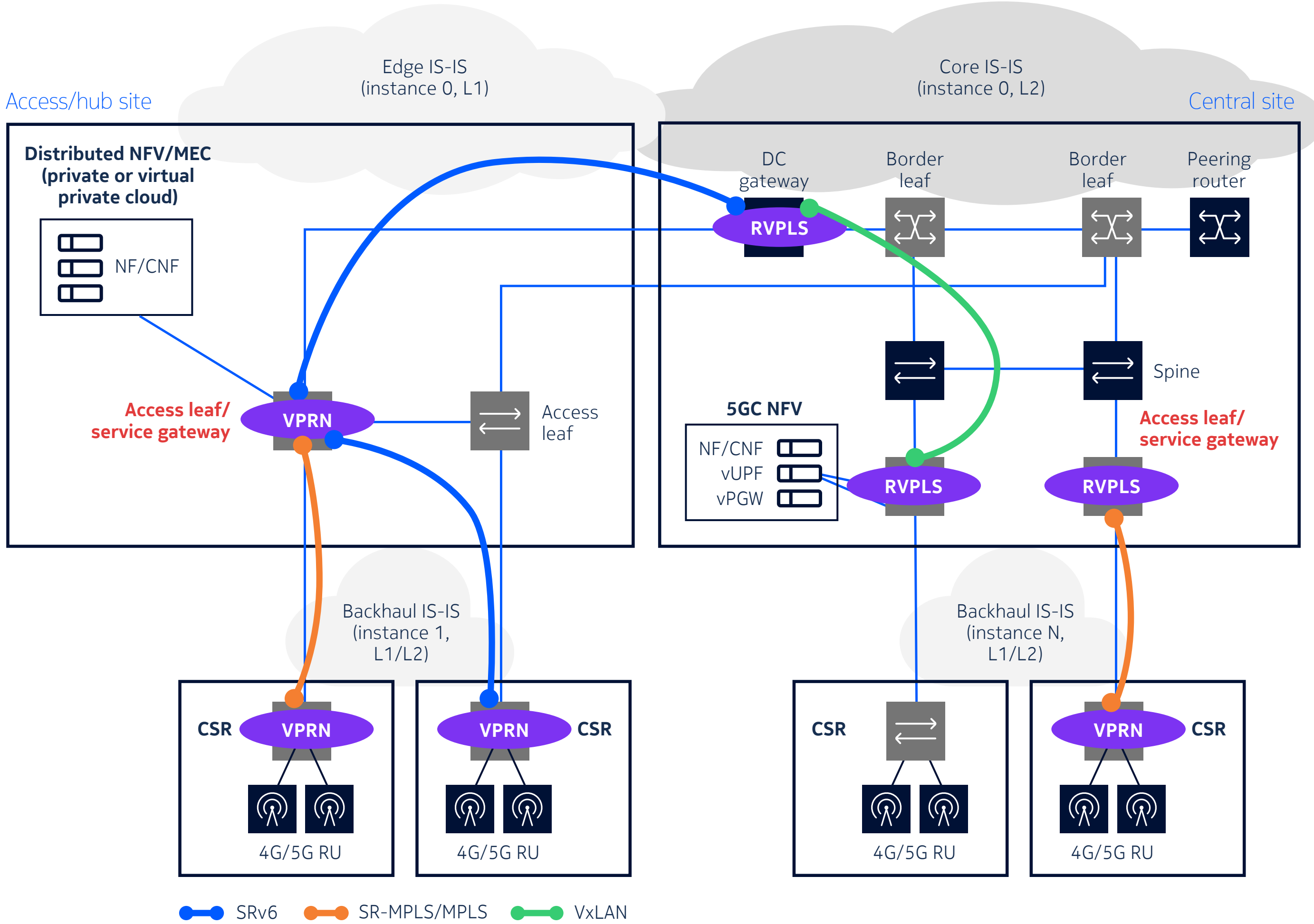
Mobile backhaul and edge networks

With the advent of 5G, many new applications are emerging for the utilities, automotive and health care industries. At the same time, the explosion of Internet of Things (IoT) devices means the number of service access points is drastically increasing.

To meet the low latency requirements of many of these applications, network operators are distributing some of the 5G core network functions closer to access points. These access points need to be dynamically connected with virtualized network functions in the distributed central offices and data centers that form the telco cloud edge and core. Interconnecting these network functions requires scalable and reliable mission-critical network services that can deliver deterministic quality of service (QoS), security and traffic isolation.

The telco cloud edge can be implemented in the operator's private network data center or as a virtual private cloud (VPC) in a public cloud. Increasingly, private clouds and VPCs are being implemented in IPv6 underlay networks. Given the large number of Kubernetes clusters that need an IP address for interconnection, an IPv6 underlay is a sensible choice. This driver, combined with the need to overlay VPN services to isolate the various services and apply different route constraints and rules, makes both SR-MPLS and SRv6 technologies of choice for the underlay and the service overlay.

Figure 6. SRv6 in the telco cloud edge



Network operators have successfully deployed SR-MPLS in the telco cloud edge within a private network data center. SRv6 is a strong alternative for implementations in private network data centers, and an even stronger alternative within the public cloud, where an IP tunneling mechanism is needed.

Another advantage of SRv6 that helps operators scale the edge network to hundreds of thousands of Kubernetes clusters is the ability to couple the overlay with the underlay. Individual services running in containers can be assigned one or more IPv6 addresses as service SIDs from the same subnet as that of the SRv6 underlay and yet can operate in their own separate overlay VPN. This requires the implementation of additional security measures to protect the SRv6 infrastructure. Such measures include access control lists (ACLs) to filter out IPv6 packets that have a destination address matching an allocated service SID, and that are received on an untrusted interface or a trusted interface with a source address outside the SRv6 domain.

SRv6 support for source routing will allow network operators to deploy more advanced applications, including:

- TE with bandwidth optimization achieved through SID compression using SRv6 (uSID)
- 5G slices with Flex-Algo TE topologies or SRv6 TE policy
- Load balancing flows (ECMP or weighted ECMP) across multiple source-routed paths
- Forwarding flows over disjointed source routed paths
- Service chaining and load balancing across multiple service functions.

These applications have limited or no inherent support from Virtual Extensible LAN (VXLAN) transport tunnels deployed in central telco cloud data centers.

Another important aspect to consider when deploying SRv6 in the telco cloud edge is to deal with the de facto segmentation of the network. Heterogeneous transport is a reality. Network operators can transition to new technologies such as SRv6 domain by domain by positioning service gateways at specific domain boundary points, as illustrated in Figure 6. Existing MPLS and new SRv6 backhaul routers must seamlessly connect to the SRv6-based edge cloud data centers and the VXLAN-based central data center.

The multi-instance service gateway, unique to Nokia routers and enabled on access leaf and data center gateways, provides key capabilities for seamless multi-domain and multi-technology communication that:

- Interworks VXLAN, MPLS and SRv6 (classic or micro-segment) at the transport and service layers
- Implements dual-instance VRF (IP-VPN to IP-VPN, IP-VPN to EVPN-IFL, EVPN-IFL to EVPN-IFL)
- Implements dual-instance layer 2 service (EVPN-VPLS or EVPN-VPWS)
- Translates between IPv4 and IPv6 BGP next hops in the base router and VRF, which enables seamless communication across the IPv4 and IPv6 transport network domains.

This multi-domain network design decouples the mobile edge and core from the backhaul, which creates significant advantages such as:

- Providing a control point for applying route constraints and policies
- Aggregating VPN routes to scale backhaul and edge/core across thousands of service endpoints (Kubernetes clusters and cell-site routers)
- Deploying SRv6 in backhaul and telco edge networks is not dependent on an upgrade of all service endpoints.

IPv6 multi-domain backbone networks

The deployment of a new backbone is a good opportunity to implement IPv6 transport in the network as IPv4 addresses become scarce. The IS-IS routing protocol is implemented with a flexible type–length–value (TLV)-based scheme that allows easier protocol extensions. This, along with its ability to run IPv4 and IPv6 concurrently within a single process, has made IS-IS the protocol of choice in many network operator backbones.

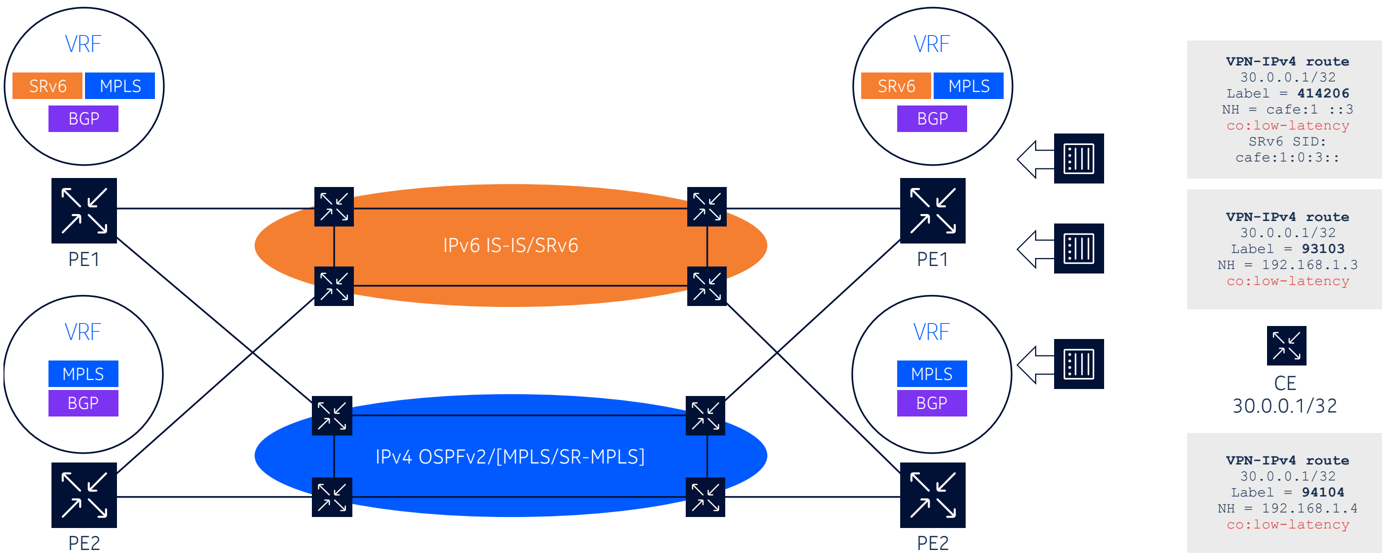
Network operators that have an IPv4 transport network based on OSPFv2 and want to migrate to a new IPv6 transport network can choose to deploy an IS-IS or OSPFv3 process. IS-IS brings all the benefits described above if the operator acquires the necessary in-house operational skills.

The implementation of an IPv6 transport network also opens the door to the introduction of SRv6 to support overlay services and TE applications. Locator summarization at domain boundaries allows a more scalable multi-domain network without propagating the system or loopback interface routes of the PE nodes between domain IGPs or into BGP-LU, as in MPLS-based networks.

The migration to IPv6 can take a considerable amount of time and is greatly assisted by the deployment of a dual-plane core backbone design. Figure 7 illustrates such a design, which allows the old OSPFv2 backbone, typically running MPLS or SR-MPLS, and the new IS-IS backbone, running SRv6, to coexist.

This application illustrates another major role of the Nokia router multi-instance service gateway: allowing the operation of dual control and data planes with the ability for traffic to fail over from one plane to the other.

Figure 7. Dual-plane SRv6/SR-MPLS backbone network



In Figure 7, an MPLS provider edge (PE) router advertises a customer edge (CE) router as a MPLS VPN route with a service label. It can only resolve an imported remote CE MPLS VPN route to an SR-OSPF tunnel.

An upgraded SRv6/MPLS PE router supporting the multi-instance service gateway advertises the MPLS and SRv6 VPN routes of a CE router . It resolves the imported MPLS and SRv6 remote CE VPN routes based on BGP best path rules and prefers the SRv6 VPN route if all other selection rules result in a tie. The user can enable Edge Prefix Independent Convergence (PIC) to use the BGP path over the MPLS/SR-MPLS backbone as a backup path for fast reroute failover of the VPN service.

Conclusions and Nokia differentiators

SRv6 is a powerful and proven technology for deploying highly programmable IP networks that meet deterministic service-level objectives for cost, performance and reliability. SRv6 addresses and mitigates most of the operational limitations and scalability issues of legacy TE and protection approaches using Label Distribution Protocol (LDP) or Resource Reservation Protocol – Traffic Engineering (RSVP-TE) and enables a wide range of new applications in the mobile edge and IPv6 backbone.

Nokia routing platforms offer a comprehensive segment routing toolkit for IP/MPLS networks, next-generation IPv6 networks and the transition between them. The segment routing and SRv6 capabilities that differentiate Nokia routing platforms include:

- Data path programmability and flexibility
  - Support of variable classic SID length (32-bit, 16-bit or 128-bit), compressed uSID and any future evolutions
  - Industry-leading Maximum SID Depth (MSD), with up to nine classic SIDs and up to 47 uSIDs, including primary path segment list, LFA backup path segment list and service SID
  - Concurrent enabling of classic SRv6, SRv6 uSID, and MPLS/SR-MPLS on the same router, service or IP interface
- Support of layer 3 and layer 2 services in the base algorithm and Flex-Algo
  - IP-VPN, EVPN-IFL, EVPN-VPLS, EVPN-VPWS
- Extensive protection and restoration capabilities
  - Base LFA, remote LFA, TI-LFA, link, node and SRLG protection
  - ECMP and primary/secondary paths
- SRv6 OAM for underlay (Locator, SID, SRv6 policy) and overlay (prefixes resolved to SRv6 tunnel)
- Scalable intra- and inter-domain SR-TE support through Locator summarization, Flex-Algo and SRv6 policy
- The industry's most flexible intent-aware segment routing gateway solution
  - Any transport to any transport: MPLS, SR-MPLS, SRv6 classic SID, SRv6 uSID, VXLAN, MPLS over UDP
  - Any service or routing context: IP-VPN, EVPN L2/L3, base router BGP routes
- Standards compliance and multivendor interoperability proven in EANTC interoperability tests and real-world production environments.

Consult our product documentation and user guides for detailed information on supported features and platforms or contact Nokia Sales to learn more.

Abbreviations

BGP	Border Gateway Protocol
BGP-LS	Border Gateway Protocol – Link State
CE	customer edge
CIDR	Classless Inter-Domain Routing
CLI	command-line interface
CNF	cloud-native network function
CSR	cell site router
EANTC	European Advanced Network Testing Center
ECMP	Equal Cost Multi-Path

EVPN	Ethernet VPN
EVPN-IFL	Ethernet VPN – Interface-less
FIB	forwarding information base
Flex-Algo	Flexible Algorithm
IGP	Interior Gateway Protocol
IoT	Internet of Things
IP	Internet Protocol
IPFIX	IP Flow Information Export
IS-IS	Intermediate System-to-Intermediate System
LDP	Label Distribution Protocol
LFA	Loop-Free Alternate
LSP	Label Switched Path
MEC	multi-access edge computing
MPLS	Multiprotocol Label Switching
MSD	Maximum SID Depth
NF	network function
NFV	network function virtualization
NSP	Network Services Platform
OAM	operations, administration and maintenance
OSPF	Open Shortest Path First
PCE	Path Computation Element
PCEP	Path Computation Element Protocol
PE	provider edge
PIC	Prefix Independent Convergence
QoS	quality of service
RLFA	Remote LFA
RSVP	Resource Reservation Protocol
RSVP-TE	Resource Reservation Protocol - Traffic Engineering
RU	radio unit
SBFD	Seamless Bidirectional Forwarding Detection
SFC	Service Function Chaining
SID	Segment Identifier
SLA	service-level agreement

SR	Service Router
SRH	Segment Routing Header
SR OS	Service Router Operating System
SRLG	Shared Risk Link Group
SR-MPLS	Segment Routing over MPLS
SR-TE	Segment Routing – Traffic Engineering
SRv6	Segment Routing over IPv6
TI-LFA	Topology-Independent LFA
TLV	type–length–value
UDP	User Datagram Protocol
uSID	SRv6 micro-segment identifier
VLAN	virtual local area network
VPC	virtual private cloud
vPGW	Virtual Packet Data Network Gateway
VPN	virtual private network
VPRN	Virtual Private Routed Network
VRF	Virtual Routing and Forwarding
vUPF	Virtual User Plane Function
VXLAN	Virtual Extensible LAN
WAN	wide area network

Standards and recommendations

RFC 4655	A Path Computation Element (PCE)-Based Architecture
RFC 5286	Basic Specification for IP Fast Reroute: Loop-Free Alternates
RFC 5440	Path Computation Element (PCE) Communication Protocol (PCEP)
RFC 6119	IPv6 Traffic Engineering in IS-IS
RFC 7490	Remote Loop-Free Alternate (LFA) Fast Reroute (RLFA)
RFC 8570	IS-IS Traffic Engineering (TE) Metric Extensions – Min./Max. Unidirectional Link Delay for metric for flex-algo, RSVP, SR-TE
RFC 8754	IPv6 Segment Routing Header (SRH)
RFC 8919	IS-IS Application-Specific Link Attributes
RFC 8986	Segment Routing over IPv6 (SRv6) Network Programming
RFC 9085	Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing

RFC 9252 BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)

RFC 9256 Segment Routing Policy Architecture

RFC 9259 Operations, Administration, and Maintenance (OAM) in Segment Routing over IPv6 (SRv6)

RFC 9352 IS-IS Extensions to Support Segment Routing over the IPv6 Data Plane

RFC 9350 IGP Flexible Algorithm

Draft standards

draft-filsfils-spring-net-pgm-extension-srv6-usid-15, Network Programming extension: SRv6 uSID instruction

draft-filsfils-spring-srv6-net-pgm-insertion-08, SRv6 NET-PGM extension: Insertion

draft-ietf-idr-bgpls-srv6-ext-14, BGP Link State Extensions for SRv6

draft-ietf-idr-segment-routing-te-policy-23, Advertising Segment Routing Policies in BGP

draft-ietf-idr-ts-flowspec-srv6-policy-03, Traffic Steering using BGP FlowSpec with SR Policy

draft-ietf-rtgwg-segment-routing-ti-lfa-11, Topology Independent Fast Reroute using Segment Routing

draft-ietf-spring-srv6-srh-compression-14, Compressed SRv6 Segment List Encoding in SRH

draft-voyer-6man-extension-header-insertion-10, Deployments With Insertion of IPv6 Segment Routing Headers

draft-trr-bess-bgp-srv6-args-02, SRv6 Argument Signaling for BGP Services

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: 890204 (March) CID213966