NOKIA

Network Functions Interconnect Architecture

A dynamic and smart network fabric for mobile broadband evolution, the Internet of Things and 5G

White paper

Communication Service Providers (CSPs) are continuously seeking ways to improve and optimize their infrastructures to cope with relentless demand for more capacity, faster service and higher efficiency. In addition, they must prepare for the introduction of 5G technology offering data speeds of 10 Gb/s and up for extreme mobile broadband as well as ultra-reliable low latency communication for massive and critical machine-type applications.

Moreover, the industry needs to migrate to a highly programmable and scalable cloud-native services architecture with its continued architecture and network functions disaggregation to enable more flexible deployment, while requiring the network functions to reliably communicate within dedicated slices over a common transport infrastructure.

This strategy white paper discusses the Nokia Network Functions Interconnect architecture (NF-IX). NF-IX is an integral part of Nokia's 5G Future X architecture and envisions a unified, smart and dynamic network fabric that is designed to enable CSPs on the path to 5G deliver programmable and cloud-optimized services with support for deterministic performance requirements. NF-IX is also applicable for other use cases such as webscale service delivery, which will be addressed in a separate paper.

NOKIA

Contents

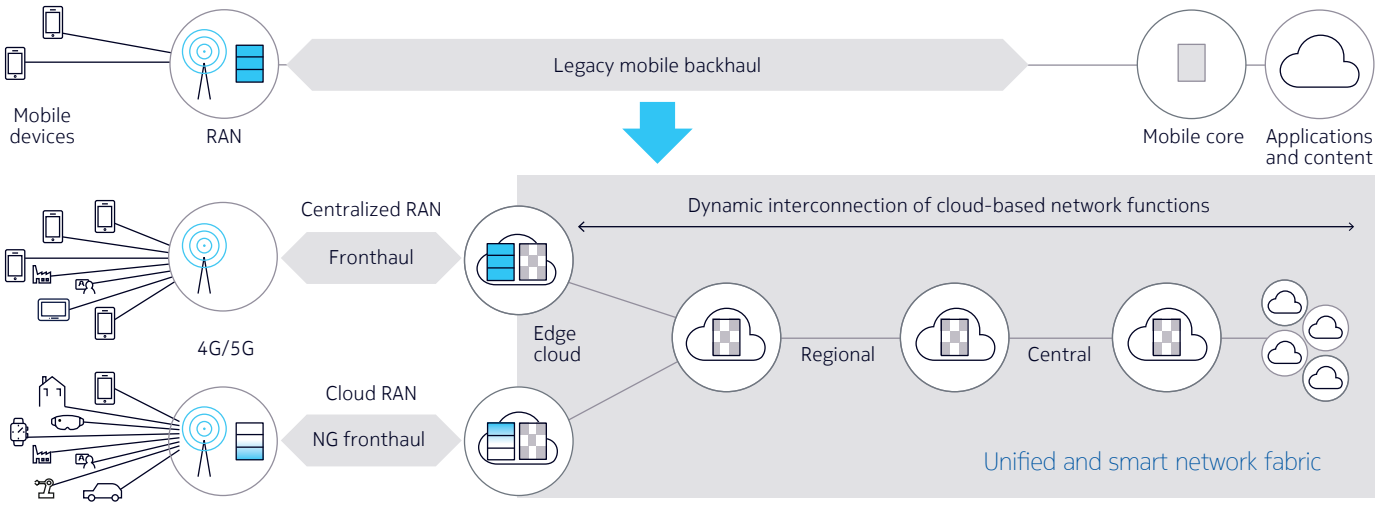
The digital network transformation to 5G

The transformative changes of the era of cloud, mobile broadband evolution and the Internet of Things (IoT) are for many network operators truly a make-or-break moment. The introduction of 5G technology will also dramatically extend the focus from consumers to industry and enterprise applications, and address a wide range of service requirements and diverse business needs:

- Extreme mobile broadband with 100 Mb/s to 10 Gb/s data rates for delivering a superior quality of experience for ultra-high definition video and immersive virtual/augmented reality
- Massive machine communications, anticipating a 10 to 100 times growth in connected devices and sensors on our person and in our homes, cars, workplaces and public infrastructure
- Critical machine communications, providing ultra-reliable, low-latency communication services for self-driving cars, remotely controlled robots and haptic feedback applications.

To deliver on these promises requires orders of magnitude of higher capacity, connectivity, agility and efficiency. The required infrastructure will necessitate a perfect union of carrier-grade IP routing and webscale IT technologies. It introduces a highly dynamic and programmable cloud-based service architecture with centralized and edge cloud infrastructures, interconnected by a unified and smart network fabric (see Figure 1). This deployment model allows a rapid introduction of new services as well as efficient and secure sharing of resources by means of network slicing techniques.

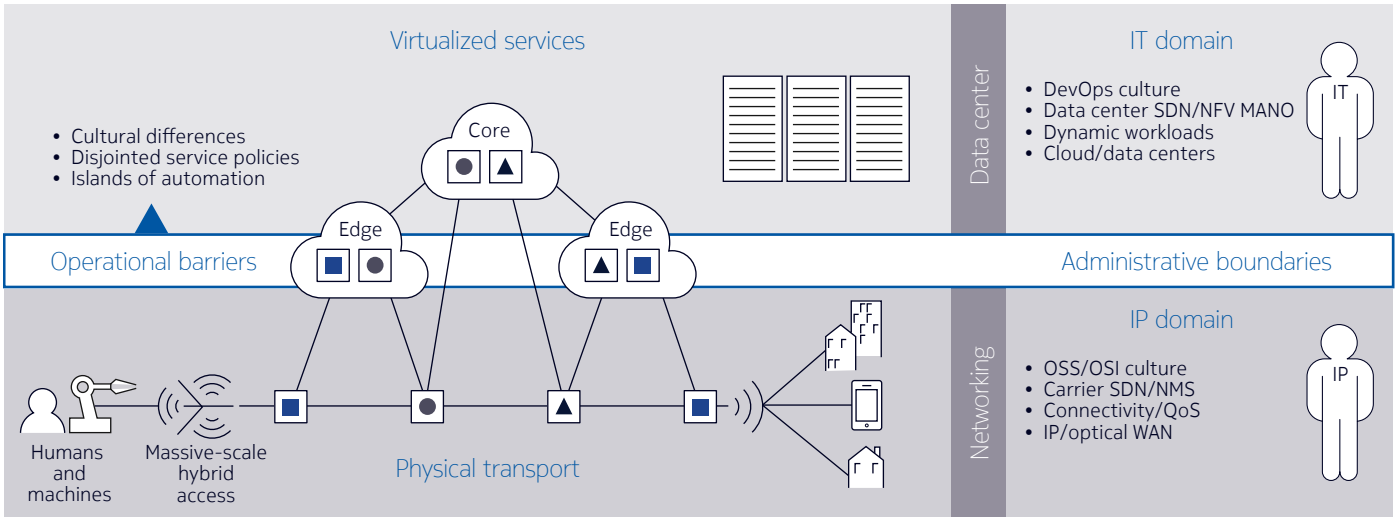
Figure 1. Evolution to a 5G service delivery architecture



As network operators adopt a DevOps culture, introduce cloud-native service models, re-architect central offices as data centers, and effectively become digital service providers, they seek to maximize IP and IT synergies. They want to consolidate network infrastructures and virtualize network functions on IT platforms without losing the inherent security and integrity of purpose-built network appliances. They need the agility and economics of webscale IT technologies while preserving the deterministic performance and mission-critical reliability of carrier-grade services. They want to run a software defined network (SDN) that can be delivered as a service and on demand.

To meet these needs, they must overcome the cultural, technological and operational barriers that historically divide the IP networking and IT domains. They must harmonize disjointed service policies, resolve a mismatch of interworking protocols, and integrate islands of automation (see Figure 2).

Figure 2. Overcoming operational barriers and maximizing IP and IT synergies



The ultimate goal is delivering virtualized services in thousands of network slices over a unified transport underlay network and automatically matching diverse application and user needs with end-to-end service delivery guarantees quickly and cost-effectively. To reach this goal requires orchestrating and interconnecting physical and virtualized network functions that are distributed over hundreds of smart central offices, regional and central data centers that collectively form the services cloud, and automatically engineering the necessary data path connectivity in the appropriate network slices. These resource needs must also be dynamically adjusted to match fluctuating service demands.

This paper introduces a solution to achieve these goals. The Nokia Network Functions Interconnect (NF-IX) Architecture is a dynamic and smart IP network fabric for automating connectivity across emerging mobile cloud service architectures with support for deterministic delivery requirements.

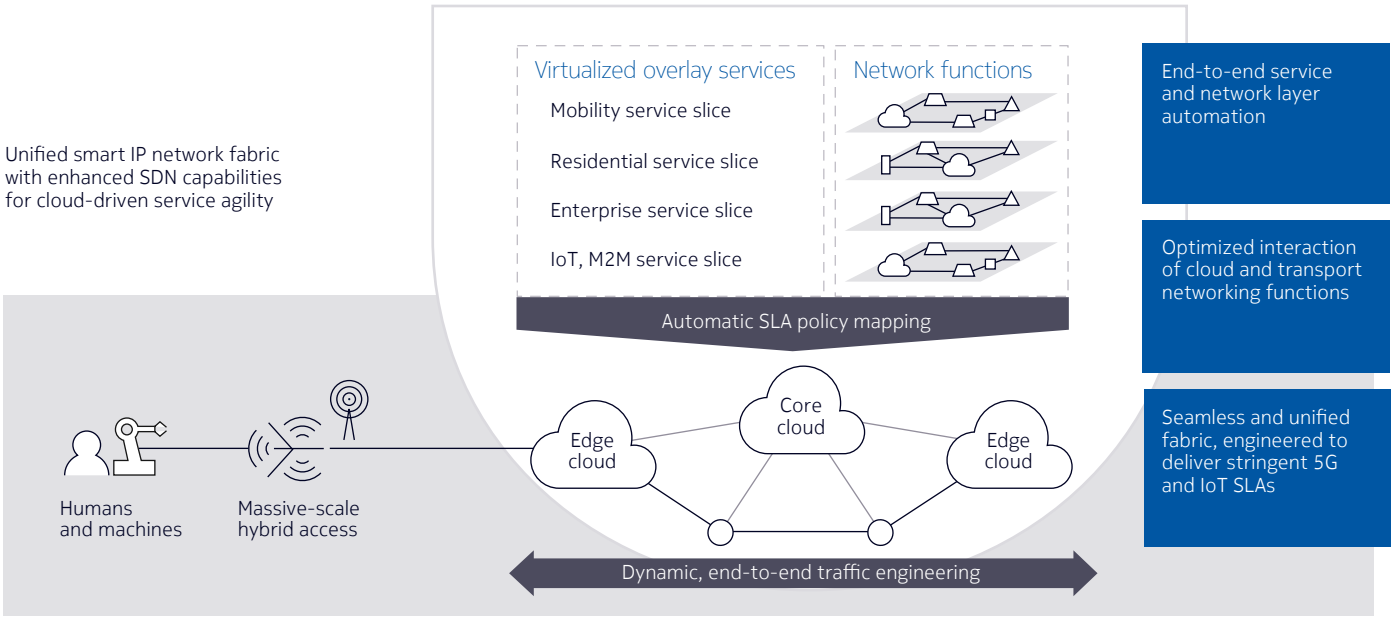
NF-IX Architecture explained

The NF-IX architecture addresses several key objectives:

- Ensures that overlay services can efficiently share IP transport resources without contention by means of a highly scalable and flexible approach to network slicing
- Maintains the functional decoupling between a cloud-native service overlay and network-native bearer services to rapidly compose new services with a broad range of delivery options
- Automatically maps overlay SLA policies on corresponding underlying transport SLAs and dynamically engineers the optimal bandwidth resources required in the WAN.

NF-IX introduces a unified and smart IP network fabric that seamlessly extends from connected clients in slices of the access network cloud to distributed network functions (NFs) in the edge cloud and core data centers (see Figure 3).

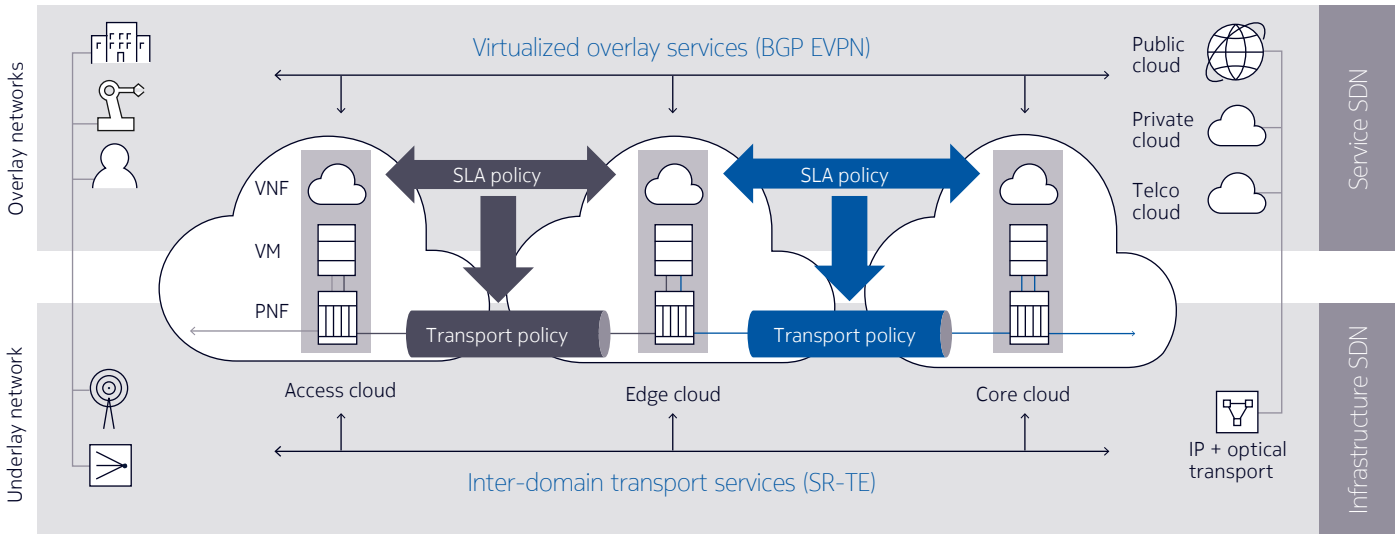
Figure 3. NF-IX architecture



While ETSI NFV MANO [1]is a very important standardization framework for building open, virtualized infrastructures, it does not define how to seamlessly interwork distributed data centres across WANs and how to orchestrate connectivity between physical network functions (PNFs) and virtualized network functions (VNFs). NF-IX addresses this gap with an open and standards-based approach that leverages proven technology building blocks available in existing network infrastructures.

NF-IX leverages the Border Gateway Protocol (BGP) and BGP MPLS-based Ethernet VPNs (RFC 7432) to acquire topology information from the virtualized service overlay and IP/MPLS WAN underlay. Segment routing is used to dynamically engineer inter-domain service tunnels between NFs that can meet deterministic bandwidth, latency and path diversity constraints end-to-end (see Figure 4).

Figure 4. Dynamic mapping of SLAs on IP/MPLS transport policies



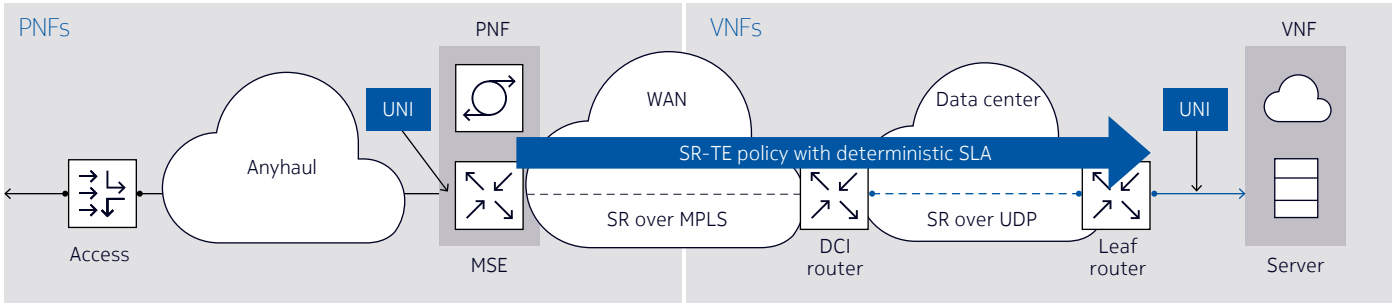
BGP Ethernet VPNs (EVPNs) [2] enable network virtualization in data centers by providing Layer 2 or Layer 3 VPN connectivity between VNFs that are part of an end-to-end service chain [3]. Network layer reachability information (NLRI) of VNFs is exchanged by means of multiprotocol extensions of the BGP control plane, which is far more scalable than conventional address flooding and learning strategies. NF-IX extends the use of EVPNs to include PNFs as well as to support services that include network-native and/or cloud-native functions.

Segment routing (SR) [4] is a highly scalable approach to source-based routing that leverages regular IGP routing protocols such as OSPF and IS-IS to distribute topology information and requires only SR head-end routers to keep forwarding state information. As a result, SR enables dynamic traffic engineering services and granular per-flow/per-application steering with various loose or strict routing constraints, including bandwidth, latency, path diversity and explicit objects to include or exclude in the route. Segment routing-traffic engineering (SR-TE) also implicitly supports equal-cost multi-path (ECMP) routing, which allows load-balancing traffic over all available links that meet the SR-TE route policy.

SR-TE policy [5]defines a traffic-engineered path between a source and a destination node through which a network function can be reached (e.g., an aggregation router, a top-of-rack (ToR) switch or a leaf router). The SR-TE policy concept decouples the forwarding control plane and data plane, which enables implementation of an abstract SR-TE policy on various underlying data path technologies, including MPLS, UDP and IPv6. This essential property enables NF-IX to provide multi-domain connectivity for network functions that may span across the IP/MPLS WAN into data center networks consisting of relatively simpler IP/Ethernet routing fabrics (see Figure 5).

An SR-TE policy is associated with a policy "color" for use by the EVPN overlay service to convey an abstract SLA requirement, and is mapped to an underlying SR-TE policy to capture the SLA semantics in an SR-TE route policy definition. How SLA policies are implemented and enforced in the underlay IP network is therefore transparent to the network functions virtualization (NFV) orchestrators composing network functions in service function chains (see Section 3) and to the VNF Managers managing the individual VNFs.

Figure 5. Extending SR-TE tunnels with end-to-end SLA policies across the WAN into the data center network

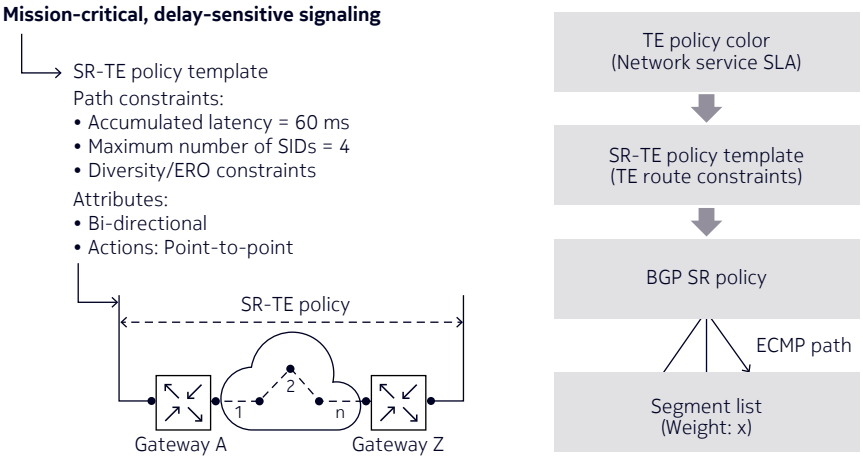


SR-TE paths implement an SR-TE policy at the data path level and are expressed as a list of segments (network prefixes, nodes or adjacencies) that the forwarding path must take to meet the SR-TE policy constraints. An SR-TE can be supported by multiple eligible SR-TE paths, in which case weights can be applied to select the most optimal path that is available.

An SR-TE-path can itself be summarized with a single segment ID (SID) or binding SID, which provides yet another layer of abstraction and scalability. The binding SID allows a leaf node to steer a traffic flow into an SR-TE policy without needing to know the forwarding details of the associated SR-TE path as it traverses the WAN domain. The binding SID also allows dynamically modifying or optimizing SR-TE paths in the WAN without impacting the SR-TE policy or the upstream nodes using the policy. SR-TE forwarding paths can be deployed on an MPLS data plane and encapsulated in UDP to cross non-MPLS domains [6].

Figure 6 illustrates the methodology by which an abstract SLA traffic policy for interconnecting NFs in the overlay service network is mapped on appropriate SR-TE paths in the underlay transport network. This approach is based on IETF recommendations and drafts that define the extensions and subsequent address family identifiers (SAFI) that enable the use of BGP to advertise and exchange SR-TE policy colors and other data across inter-domain boundaries [5] [7].

Figure 6. Dynamic mapping of SLA policies to SR-TE paths



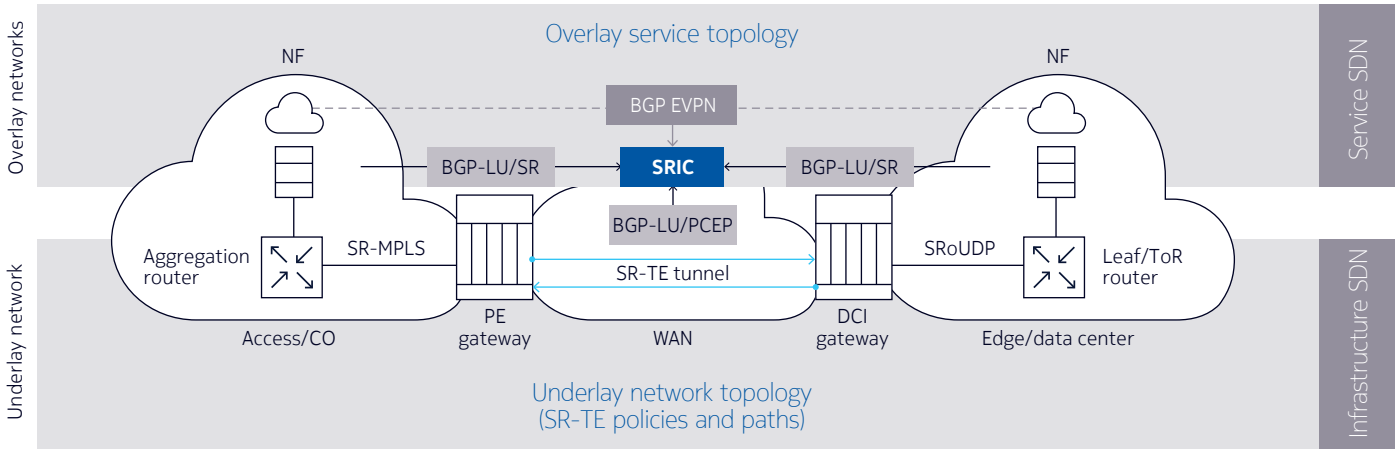
NF-IX introduces a **Segment Routing Interconnect Controller** (SRIC) to automate the process of mapping SLA requirements on SR-TE policies and subsequent SR-TE paths across the WAN.

The SRIC performs the following key tasks:

- Defines the catalog of all SLA policy colors and their associated SR-TE policy templates
- Discovers EVPN routes and maintains a virtual service connectivity map of all NFs
- Discovers WAN routes and reachability of all connected data center gateways
- Creates and/or discovers available SR-TE paths for resolving SR-TE policy requests
- Collects link statistics through telemetry interfaces to assure SLA policy goals.

The SRIC makes extensive use of the BGP to discover [7] network layer reachability information across WAN boundaries and between data centers of both the overlay and underlay network layers (see Figure 7). MP-BGP EVPN extensions are used to exchange reachability information of network functions and EVPNs. SR-TE extensions to BGP-LU (Labeled Unicast) and BGP-LS (Link State) [8][9] are used to exchange SR-TE topology information. BGP communication may be facilitated by a BGP route reflector hierarchy, which is usually the default in data center networks, but may also occur directly through BGP peering with DCI routers at the WAN demarcation. Other protocols, such as NETCONF/YANG and PCEP (RFC 5440) with extensions for segment routing [10], can also be used to acquire topology information.

Figure 7. Segment Routing Interconnect Controller



SR-TE policy auto-instantiation can occur automatically upon receiving a BGP update on a new EVPN route, which includes the endpoints (BGP next hop) of the network function instances to be interconnected, the route target and policy color. The new EVPN route may, for example, be originated by the data center SDN controller(s) upon VNF instantiation or by a virtualized infrastructure manager that is part of an NFV MANO environment. EVPN reachability information is exchanged between the BGP processes managing the near- and far-end domains in which the NFs reside, after which the EVPN routes are resolved and advertised by means of BGP-LU updates between the SRIC and DCI routers with segment routing capability.

After the SRIC has received the new EVPN routes as well as BGP-LU routes for the aggregation/leaf routers and DCI routers, it can build a service connectivity map in overlay with the IP transport topology. It will then look up the requested SLA policy color for the EVPN route and apply the corresponding SR-TE policy template to calculate an optimal SR-TE path through the underlay transport network. In the simplest case, the SRIC creates two symmetric unidirectional tunnels based on receiving two EVPN routes with the same color and route target. The SRIC concludes the process by selecting and binding the SR-TE path to the SR-TE policy and signaling the binding SID through BGP to enable the SR-TE policy installation on the NF leaf routers.

When a network service is terminated and EVPN routes are withdrawn, the supported SR-TE policies are automatically removed as well.

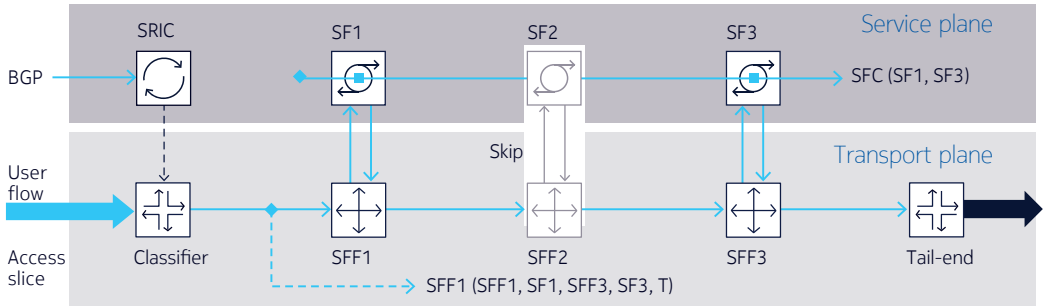
Service function chaining

Service function chaining (RFC 7665) is a powerful concept for composing programmable and flexible end-to-end services. A Service Function Chain (SFC) is created by concatenating service functions (SFs) in a specific order and subsequently steering user flows along the resulting Service Function Path (SFP). SFs are logical functions (e.g., a firewall) that can be physical or virtualized and that operate on various layers of the OSI stack. An SFC can perform various end-to-end services by including or excluding specific SFs in the SFC transport path. SFCs can be quite complex, with various branches, loops or jumps. The SFs in an SFP may reside on different servers data centers or network devices, and a key task of the NF-IX smart fabric is to provide location transparency.

At the head-end of an SFC is a Classifier function that matches incoming user flows against an SFC policy to select the applicable SFP that renders the desired end-to-end service (see Figure 8). The Classifier then encodes the SFP forwarding instructions in a Network Service Header (NSH) and composes the

corresponding SFC transport path as an ordered list of service segments that must be traversed along the service chain. An SFP doesn't necessarily include all SFs (e.g., SF2 in Figure 8).

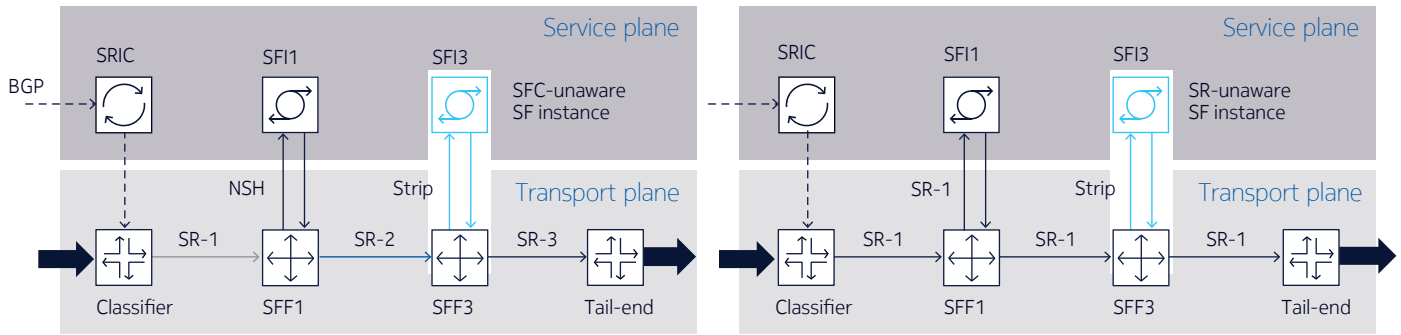
Figure 8. Service Function Chaining architecture



Each SF is associated with a Service Function Forwarder (SFF) that processes the SFP header information; forwards the user packet to the SF performing the service; updates the NSH information of the processed packet; and passes it on to the next SFF in the SFP. SF instances may be integrated with a specific SFF on the same physical device or server or located on a different device or server. An SFF may act on behalf of multiple SF types and/or instances and enable dynamic load-balancing. SFFs and SFs are considered network functions from an NF-IX perspective once the SRIC learns their network layer reachability information through BGP updates.

The SFC service plane is separated from the transport plane to allow for a flexible mapping on various data path implementations, such as SR-MPLS, SR-MPLS over UDP and SRv6. SFC NLRI is exchanged through BGP control plane updates using new address family extensions for NSH and segment routing-based service chaining [11] [12]. These BGP control plane extensions enable the SRIC to learn the SFC topology and construct the various SFPs by including and concatenating service segments in the order of their advertised Service Index (SI).

Figure 9. SFC transport path implementation models



Decoupled service function path (NSH chaining)

Coupled service function path (SR chaining)

The SFC transport plane can leverage segment routing [13] using a decoupled or coupled SFP implementation model [14]:

- Decoupled (NSH-based SFC with SR-based transport tunnels): Segment routing provides the individual transport tunnels between SFFs, and NSH is used for service function chaining and associated meta-data (see the left of Figure 9).

- Coupled (SR-based SFC with integrated NSH service plane): Each service hop is represented as a segment of the SR segment list. Segment routing is responsible for steering traffic through all necessary SFFs in the segment routing path. NSH is responsible for maintaining the service plane and holding the SFC instance context and metadata (see the right of Figure 9).

Some SFs may be SFC-unaware or incapable of processing NSH and/or SR header data. In that case, the associated SFFs must strip the SFC header data from the packet prior to forwarding it to the SF, and re-impose it again on the returned traffic. Because market adoption of NSH-aware SFs is currently more advanced, the decoupled model is initially more widely applicable than the coupled model until SR-aware SFs become mainstream.

Using either implementation scenario, the SFFs of each service segment are interconnected with SR-TE transport tunnels, and the advertised Service Path Identifier (SPI) is assigned. End-to-end services can subsequently be performed by dynamically selecting SFPs based on the policy decisions made by the classifier at the SFC head-end.

Network slicing

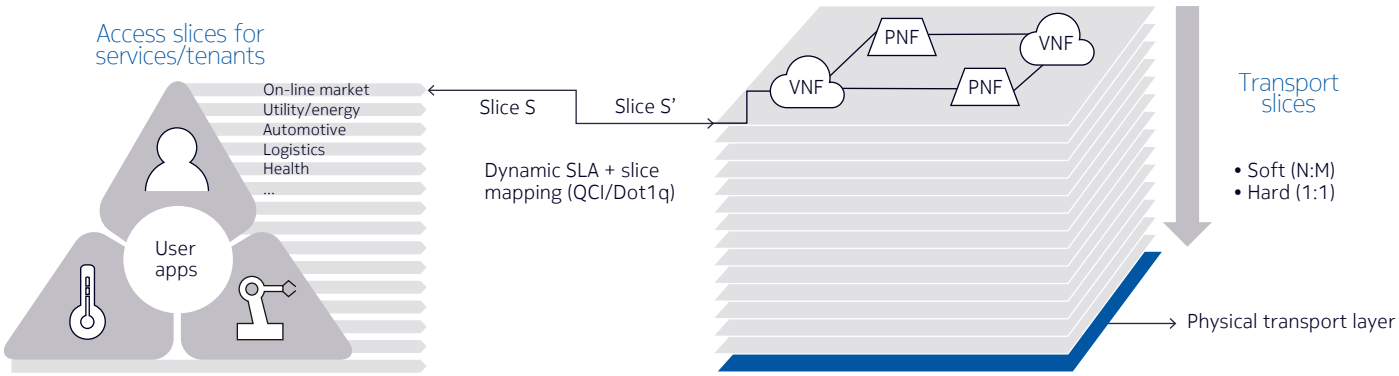
Network slicing will be an essential feature for mobile cloud infrastructures that allows a multitude of services with different requirements to share a common physical bearer network without resource contention issues. A network slice is a logical or virtual network partition with dedicated resources to support an isolated set of services, applications or users/tenants.

Network slicing typically starts in the multi-access network where traffic flows are classified and marked with a specific Class of Service (CoS) profile (e.g., 3GPP QoS Class Identifier, IP DSCP or 802.1q VLAN tag). The traffic flows are subsequently mapped into a proper transport network slice that supports the end-to-end transport SLA that meets the latency, delay, jitter and packet loss objectives of the CoS profile.

A distinction is made between “hard” and “soft” slicing. Hard slicing is required for critical machine-type and control applications and provides bandwidth and latency guarantees, even under failure conditions. Soft slicing allows for a controlled overbooking of transport resources, which results in a more economical use of network resources for high-volume applications with relaxed latency constraints, such as streaming video.

To set up the NF-IX smart fabric for network slicing, appropriate policy colors are defined for each CoS profile, with corresponding SR-TE policy templates that stipulate bandwidth, latency, SRLG path diversity and other SLA constraints to be met by the bearer transport service (see Figure 10).

Figure 10. Mapping access network slices on NF-IX transport slices



Soft slicing is the default option for transport services and typically results in SR-TE tunnels that are carried over a set of ECMP-aware SR paths with N+1 redundancy and controlled overbooking. Hierarchical QoS may be applied to service flows on SR-TE tunnel ingress when multiplexing multiple service flows into a specific network slice. Hard slicing results in the SRIC provisioning the SR-TE tunnel with dedicated physical transport resources for its exclusive use to ensure commited information rates and deterministic latency characteristics.

The SRIC ensures that agregate service flows do not exceed the nomimal SR-TE tunnel capacity and also guarantees restoration in case of failure. Due to node or link failures, maintenance activities or network congestion, certain traffic engineering characteristics of the underlying IP transport network may change over time that potentially impact the SLA conformance of overlay sevices. To ensure SLA conformance, the SRIC collects real-time traffic engineering metrics, conducts periodic latency measurements of SR-TE tunnels in service, and dynamically re-sizes or re-optimizes SR paths as needed.

After proper SLA policy colors and SR-TE policy templates are defined, the NF-IX smart fabric can dynamically classify and steer ingress service flows from access slices into appropriately engineered soft or hard SR-TE service tunnels and re-mark packets with a corresponding CoS (e.g., Diffserv codepoint) in the IP bearer network that reflects the application's QoS objectives.

Deployment assumptions

Nokia NF-IX leverages existing deployments and capabilities to minimize cost, risk and time.

The **data center network fabric** is assumed to be IP/Ethernet only with support for ECMP using Layer 3 and Layer 4 packet header information. MPLS support is optional, and BGP is the only routing protocol used. Each network device in the data center network is assumed to peer with all its connected neighbors.

The **WAN** is required to support SR-TE in combination with an MPLS data plane and a full set of IGP/BGP routing protocols. DCI routers are the demarcation devices between data centers and the WAN. The data centers and WAN are assumed to be in different administrative domains (i.e., they have different BGP Autonomous System Numbers).

Network function connectivity must be supported for the following options:

- Directly on the IP underlay: SR policy endpoint is VNF
- On top of EVPN signaled overlay: SR policy endpoint is a leaf router/ToR switch or a vSwitch
- PNF connected to IP underlay: Network access point is an IP interface

Network virtualization overlays are assumed to be based on BGP MPLS-based Ethernet VPNs.

NFV service orchestration is assumed to be present and to conform to ETSI NFV MANO standards.

Access networks, data centers and the WAN are assumed to be managed by dedicated SDN domain controllers that may have been developed in-house or sourced from various vendors.

NOKIA

Conclusion

Nokia NF-IX is an elegant and innovative approach to help CSPs make an agile, high performance, reliable mobile oriented cloud services architecture a practical reality and a commercial success. With NF-IX they can realize their goal of becoming digital service providers by transforming and unifying their networks into agile, scalable and highly automated multi-cloud infrastructures. Without NF-IX, IT has to manually interact with networking groups to orchestrate service operations end-to-end; custom tools have to be developed to execute and track network interactions between distributed data centers and the wide area network; and assigning dynamic QoS for software-defined services becomes next to impossible.

With NF-IX, they can:

- Rapidly create and efficiently deliver cloud-optimized services over a unified transport underlay by using dynamically connected service function chains to address a wide range of user needs
- Automatically map connectivity SLA requirements to underlying traffic engineering policies
- Leverage open standards and proven technologies to minimize risk, cost and time-to-market

NF-IX establishes a unified IP network fabric to serve as a common, carrier-grade transport network for a disaggregated and cloud-native service overlay. The adoption of network slicing techniques allows a multitude of services with different SLA requirements to efficiently share this common transport underlay without resource contention issues. Various SR-TE extensions to the BGP control plane enable NF-IX to dynamically engineer scalable and granular service tunnels that support these SLA requirements. These service tunnels are used to automatically interconnect network functions into service chains that deliver the end-user experience.

NF-IX is deployable on existing data center and WAN infrastructures and allows the underlying data forwarding plane to evolve with minimal impact on the services plane. Operators can start with a simple IP/Ethernet data center fabric (i.e., segment routing over UDP) and optimize if native MPLS switching is available (i.e., segment routing over MPLS) without requiring Day One changes to IP VPN and EVPN service signaling.

NF-IX leverages the learnings and best practices from the Nuage Networks Virtualized Services Platform (VSP) for data center automation and the Nokia Network Services Platform (NSP) for carrier SDN service automation and network optimization.

Nokia is actively engaging with industry partners and standardization forums to elaborate and validate the NF-IX solution approach and enhance open-source code with the required programmatic interfaces for NF-IX.

For more information about Nokia innovations for mobile broadband evolution, IoT and 5G, visit the Nokia Networks 5G web page, cloud native and IoT web pages.

NOKIA

References and resources

1. Network Functions Virtualization. Architectural framework. ETSI GS NFV 002
2. BGP MPLS-Based Ethernet VPN. RFC 7432
3. Service function chaining (SFC) architecture. RFC 7665
4. Segment routing architecture. Draft-spring-segment-routing
5. Segment Routing Policy for Traffic Engineering. Draft-spring-segment-routing-policy
6. Encapsulating MPLS in UDP. RFC 7510
7. Advertising Segment Routing Policies in BGP. Draft-idr-segment-routing-te-policy
8. BGP link state extensions for segment routing. Draft-idr-bgp-ls-segment-routing-ext
9. Northbound distribution of link state and traffic engineering information using BGP. RFC 7752
10. PCEP extensions for segment routing. Draft-ietf-pce-segment-routing
11. BGP control plane for NSH SFC. Draft-bess-nsh-bgp-control-plane
12. BGP extensions for Segment Routing based SFC. Draft-dawra-idr-bgp-sr-service-chaining
13. Segment Routing for Service Chaining. Draft-xuclad-spring-sr-service-chaining
14. NSH and Segment Routing Integration for Service Function Chaining. Draft-guichard-sfc-nsh-sr
15. MPLS segment routing in IP networks. Draft-bryant-mpls-unified-IP-SR

NOKIA

Abbreviations

BBU	baseband unit
BGP	Border Gateway Protocol
BGP-LS	BGP-Link State
BGP-LU	BGP Labelled Unicast
CoS	Class of Service
CPC	Cloud Packet Core
DCI	data center interconnect
DSCP	Differentiated Services Code Point
ECMP	Equal-cost multi-path routing
ETSI	European Telecommunications Standards Institute
EPC	Evolved Packet Core
ERO	Explicit Route Object
EVPN	Ethernet Virtual Private Network
GW	gateway
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IoT	Internet of Things
IS-IS	Intermediate System to Intermediate System
M2M	machine-to-machine
MANO	management and orchestration
MPLS	Multiprotocol Label Switching
MSE	multi-service edge
NF-IX	Network Functions Interconnect
NFV	network functions virtualization
NG	next-generation
NLRI	network layer reachability information
NSH	Network Service Header
NSP	Network Services Platform
NVO	Network Virtualization Overlay
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
OSS	operations support system
P2P	point-to-point
PE	provider edge
PCEP	Path Computation Element Protocol
PNF	physical network function
QCI	Quality of Service class identifier
QoS	Quality of Service
RAN	radio access network
SDN	software defined network
SF	service function
SFC	Service Function Chain
SFF	Service Function Forwarder
SFI	Service Function Identifier
SFP	Service Function Path
SI	Service Index
SID	segment ID
SLA	Service Level Agreement
SR	segment routing
SRIC	Segment Routing Interconnect Controller
SR-TE	segment routing - traffic engineering
TCO	total cost of ownership
ToR	top-of-rack
UDP	User Datagram Protocol
UNI	user-network interface
VLAN	virtual local area network
VNF	virtualized network function
VPN	virtual private network
VSP	Virtualized Services Platform
WAN	wide area network

About Nokia

We create the technology to connect the world. Powered by the research and innovation of Nokia Bell Labs, we serve communications service providers, governments, large enterprises and consumers, with the industry's most complete, end-to-end portfolio of products, services and licensing.

From the enabling infrastructure for 5G and the Internet of Things, to emerging applications in digital health, we are shaping the future of technology to transform the human experience. networks.nokia.com

Nokia Oyj
Karaportti 3
FI-02610 Espoo, Finland
Tel. +358 (0) 10 44 88 000

Document code: SR1803023133EN (April)