Flexible Service Chaining
## Contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive summary</td>
<td>03</td>
</tr>
<tr>
<td>Main driver for Flexible Service Chaining</td>
<td>04</td>
</tr>
<tr>
<td>Flexible Service Chaining</td>
<td>06</td>
</tr>
<tr>
<td>Architecture options for Flexible Service Chaining</td>
<td>07</td>
</tr>
<tr>
<td>3GPP flexible mobile service steering</td>
<td>07</td>
</tr>
<tr>
<td>IETF and ONF Service Function Chaining</td>
<td>11</td>
</tr>
<tr>
<td>Service Chaining in a Telco cloud</td>
<td>12</td>
</tr>
<tr>
<td>Nokia Service Chaining solution target architecture</td>
<td>12</td>
</tr>
<tr>
<td>Conclusion</td>
<td>14</td>
</tr>
<tr>
<td>References</td>
<td>15</td>
</tr>
</tbody>
</table>
Executive summary

Today’s networks need to serve an increasing number of services and applications with different Quality of Experience (QoE), security and charging requirements. Service Chaining defines how network services for applications such as watching a video, browsing the web, playing online games or using popular messaging and chat applications are implemented through the treatment of packets by a set of service functions.

A Service Chain defines the service functions (e.g. Firewall, Deep Packet Inspection or Network Address Translation) and their order that must be applied to packets and/or frames in the downlink and uplink path. A Service Function Path is a forwarding path used for delivery of the traffic along a Service Chain. A Service Function Path may be uni-directional or bi-directional and is built from a series of Service Functions and appropriate forwarding elements. Service Function Paths determine explicit instances of service functions and forwarding elements to be used for specific traffic flows.

Service Chains already exist in mobile networks: DPI and Firewalls for Internet access and NAT for access to operator sub-networks using private IP addresses. Service Chains are typically invoked based on Access Point Names (APN). The APN maps to the Packet Data Network (PDN), which indicates the type of network access (e.g. Internet access, access to a corporate network or an operator sub-network). This information is static for all users and is a limiting factor in today’s deployments. Operators would benefit from the ability to invoke Service Chains per subscriber, per application and on demand, rather than steering all traffic towards a specific APN and Service Chain.

Flexible Service Chaining is the ability to choose the relevant Service Chain through the analysis of a traffic flow, subscriber information and the application used. Flexible Service Chaining negates the need of statically configuring Service Chains per APN, as flow classification applies to a specific Service Chain. Flexible Service Chaining also offers additional opportunities to monetize services for operators, through the control and enrichment of certain services, such as online gaming or Voice over IP. Traffic steering through Service Functions that are used to allow or disallow and/or improve services for a certain user (e.g. DPI or content filtering) could be based on subscription information and the application type. This added flexibility will invoke complex and costly Service Functions (e.g. DPI) only when needed. However, Flexible Service Chain setup also requires sophisticated control of the Service Functions and the Service Function Paths in the network.

Service classification is an important aspect for Flexible Service Chaining. Classification up to Layer 4 is sometimes necessary for Service Chains, whereas classification up to Layer 7 may be needed for other Service Chains. Classification for both uplink and downlink traffic flows delivers the most benefit. Classification may be needed at the ingress and the egress of the operator’s packet network, and is recommended to be performed every time a flow leaves a Service Function and before it is forwarded to the next one. For this reason, Service Functions are not
recommended to directly communicate with each other, but instead use a Service Function Forwarder (SFF) entity. Each time an SFF receives a flow it performs classification up to Layer 4, and based on the results of the classification and its forwarding rules decides how to further steer the traffic. If an SFF is not capable of establishing how to steer the traffic by Layer 4 classification, it forwards the traffic to a Layer 7 classifier to perform the task. Classification at each stage enables an advanced scenario of a Dynamic Service Chain, where at runtime as the flow traverses the Service Chain, the decision is made whether to access all of the Service Functions or only a specific subset. Dynamic Service Chains is another useful tool for operators to optimize usage of their network resources, as Service Chains are only invoked on demand depending on classification results but the contents of each chain can be further optimized based on subsequent classification results.

Closer to today’s static service chains is the approach of classifying only at the ingress and the egress of the operator’s SGi LAN, and based on this classification ‘paint’ packets with a specific color. This is a step forward to the currently applied service chains but does not support Dynamic Service Chains and limits flexibility, especially in cloud environments, where Virtual Network Functions (VNFs) scale and are self-healing.

Flexible Service Chaining becomes totally flexible when adopting an architecture that is implemented within a Telco cloud, based on cloud techniques and Software Defined Networking (SDN).

Main driver for Flexible Service Chaining

Operators provide Value Added Services (e.g. parental control, firewall or video optimization) in a data network by chaining Service Functions within the data path, and by forwarding traffic through each Service Function (also referred to as Service Enabler) in a certain sequence. Mobile operators deploy Value Added Services networks, also referred to as the SGi LAN (SGi is the interface at the Packet Data Network Gateway (P-GW) towards Packet Data Networks) in order to improve the user’s QoE and to provide Value Added Services to certain groups of users, as illustrated in Figure 1. Service Functions in this context can also be used by the operator to control the user’s traffic (e.g. by DPI) and take appropriate actions, for example, to downgrade or upgrade Quality of Service (QoS), block traffic dependent on user profile, network operator’s policies, Radio Access Technology (RAT) type or application characteristics.
Two types of Service Chain models are used in today’s networks. The static serial model is where Service Enablers are statically connected with each other, with traffic passing through all the enablers deployed in the chain. The APN is typically used to differentiate the different service chains (e.g. one for Internet access, another one for Voice over LTE, etc).

Some drawbacks of the static serial model are as follows:

- Not all Service Enablers are needed in all cases, which can add to unwanted delays and leads to inefficient usage of resources
- All enablers within the chain need the ability to handle the maximum traffic load
- High levels of OPEX are generated as it is difficult to adjust the Service Chain

The second Service Chain model is the hairpin model. All Service Enablers are connected to a centralized traffic routing controller, with traffic always routed back to the centralized controller to decide the next stage after it passes through a certain Service Enabler. An additional drawback of the hairpin model is high CAPEX, as all traffic needs to go through the centralized controller repeatedly in order to determine the next stage. This requires traffic steering and huge network capacity at the centralized controller as traffic traverses it multiple times.

Both currently used Service Chain models have drawbacks, such as rigidity and complexity and are static or at least semi-static. Given the rise in data service usage and the corresponding data traffic growth in networks, efficient management of Service Functions and infrastructure capacity becomes essential. There is a strong need for Flexible Service Chaining solutions offering the following abilities:

- A common architecture which fits to both fixed and mobile networks
- Flexible and dynamic creation, modification and deletion of Service Chains and their components

Figure 1: APN-based Service Chaining model
• Radio Access Technology agnostic
• Support of virtualized Service Functions that are elastic in terms of scalability, can be rapidly deployed for increased agility and provide topology flexibility
• Ability to support uni-directional and bi-directional Service Chains

A uni-directional Service Chain requires that traffic must be forwarded through the ordered Service Functions in one direction, whereas a bi-directional Service Chain requires a symmetric path in which the Service Function instances are the same in both directions. A hybrid Service Chain has attributes of both un-directional and bi-directional Service Chains.

Flexible Service Chaining

As illustrated in Figure 2, the definition of a generalized Service Chaining architecture is as follows:

• Network Service - an externally visible service (e.g. Voice over LTE) offered by a network operator that is built out of several Service Functions and executed in pre-determined sequences
• Service Function - a Service Function (e.g. NAT, Firewall and Deep Packet Inspection) applies a certain packet treatment, which is used independently or in collaboration with other Service Functions to enable a specific network service behavior
• Service Classifier - a function that matches traffic flows to forwarding rules that are generated based on policies, that takes into consideration customer, network and service profiles to determine and impose the Service Chain
• Service Chain - a Service Chain defines the services required, and the order that must be applied to packets and/or frames. Service Chains can be uni-directional or bi-directional, need not be linear and can be represented by a graph topology
• Service Function Path - a forwarding path used for delivery of the traffic along a Service Chain

Flexible Service Chaining can optimize the use of Service Functions in the SGi LAN by selectively steering traffic through certain Service Functions depending on a variety of factors, such as the type of service. Certain Service Functions can be bypassed completely, which helps to reduce or even avoid over-dimensioning of network resources. Efficient and flexible mobile service steering relies on a user’s subscription profile, the used access technology, application characteristics or other criteria. That information is used by the operator to define traffic steering policies that route a user’s traffic via the appropriate Service Functions.
In order to overcome the drawbacks of current Service Chaining mechanisms, new routing mechanisms are necessary. The Internet Engineering Task Force (IETF) is working on how to steer traffic in the SGi LAN domain. The relevant activity is known as Service Function Chaining (SFC) \(^2\). The IETF introduces an on-demand routing mechanism called the Network Service Header (NSH) that realizes efficient and flexible routing. The IETF’s work is aligned with general SDN principles of decoupling the control and user plane and dynamically programming the user plane by the control plane.

The Open Network Foundation (ONF) \(^3\) stems from the SGi LAN architecture of the IETF but focuses on the Open Flow (OF) protocol-based splitting and related extensions. The European Telecommunications Standards Institute (ETSI) Network Function Virtualization (NFV) with the evolving specification known as EVE describes models on how Service Chaining architecture fits to the overall NFV architecture. The next section provides an overview of all these activities.

**Architecture options for Flexible Service Chaining**

**3GPP flexible mobile service steering**

It is commonly accepted in 3GPP that the Policy Control and Rules Function (PCRF) will remain as the single policy decision point in the 3GPP architecture. PCRF functionality will be enhanced to generate traffic steering policies. For the enforcement point, 3GPP agreed the following three alternatives selected from 3GPP TR 23.718 \(^1\):

1. Leverage the existing Policy Charging and Control (PCC) framework with the Sd interface

This architecture leverages on the Traffic Detection Function (TDF) in the PCC framework to act as the policy enforcement point for traffic steering policies. TDF may support multiple instances/processes for uplink and downlink traffic (i.e. one process for uplink traffic and one
process for downlink traffic). The traffic steering policies are transferred via the existing Sd interface and can be applied on the session, application and service data flow level. The way traffic is steered in the SGi LAN is not in scope of 3GPP, with 3GPP referring this aspect to other standardization bodies. An example of traffic steering enforcement mechanisms is the utilization of the NSH header in the user plane IP packets as proposed by the IETF. This architecture is illustrated in Figure 3.

![Figure 3: Standalone TDF as policy enforcement point of traffic steering policies](image)

2. **Leverage the existing PCC framework with the Gx interface**

   This architecture option leverages on the Policy and Charging Enforcement Function (PCEF) in the PCC framework to act as the policy enforcement point for traffic steering policies. The PCEF may support multiple processes for uplink and downlink traffic (i.e. one process for uplink traffic and one process for downlink traffic). The traffic steering policies are transferred via the existing Gx interface and can be applied on the session, application and service data flow level. The way traffic is steered in the SGi LAN is again not in scope of 3GPP. Similar to the option above, traffic steering enforcement mechanisms could utilize the NSH header in the user plane IP packets as proposed by the IETF. This architecture is illustrated in Figure 4.
3. Introducing the Traffic Steering Support Function (TSSF) for traffic steering

This architecture option introduces a new functional entity TSSF for processing traffic steering rules. This requires the introduction of a new reference point St to provide traffic steering policies from the PCRF to the TSSF.

The functionality of the TSSF is as follows:

- To map traffic steering policies received from the PCRF to configured Service Chains in the TSSF
- To construct specific instances of traffic steering enforcement information of the applicable Service Chain for each relevant flow, bearer and application
- To pass traffic steering enforcement information in the SGi LAN for enforcement on user plane packets
- To receive metadata from the PCRF and pass to the entities that need it in the SGi LAN (could be both Service Functions and forwarding elements). Metadata handling and relevant reporting is not specified in 3GPP Release 13 but planned for 3GPP Release 14

The details on how the enforcement of traffic steering policy rules are performed within the TSSF is out of the 3GPP scope. The architecture illustrated in Figure 5 shows this option with the St interface and TSSF. Any interfaces inside the SGi LAN domain are out of scope of the 3GPP.
Figure 5: Reference architecture with TSSF and traffic steering policy interface St
The IETF proposes the following architecture as illustrated in Figure 6 related to Service Function Chaining:

The IETF architecture is based on a split control plane and user plane architecture and is aligned with the principles of SDN. The control plane receives traffic steering policies that take into consideration the topology of the SGi LAN domain, as well as information from the user plane (e.g. load of elements, status, etc) and constructs the Service Function paths. They are translated into forwarding rules, which are transferred by the control plane to the user plane and specifically to the classifier nodes and the SFFs. The classifier nodes are capable of performing classification using information up to Layer 7, whereas the SFFs have classification capabilities up to Layer 4. Both classifier nodes and SFFs are capable of handling the NSH \(^4\). The NSH also provides a mechanism for metadata exchange along a Service Function path. The IETF proposes Representational State Transfer (REST) based Cx (e.g. C1, C2) interfaces.

The Open Network Foundation (ONF) proposes an architecture similar to the IETF which is more focused on using OF as southbound protocol. ONF clearly describes the current capabilities of the OF to control Service Chains and any extensions needed.
Service Chaining in a Telco cloud

With the introduction of virtualization and cloud technology in mobile operator networks, physical functions will become virtual or as VNFs (as referred to by ETSI NFV). ETSI NFV has specified the relevant overall architecture. Each VNF is accompanied by its VNF Manager (VNFM), which orchestrates its lifecycle operations (i.e. instantiation, scaling, upgrades, etc). End-to-end services, for example VoLTE, require more than one virtual and or physical function interconnected with virtual links. The VNFM of each VNF does not have the visibility of the end-to-end service but is offered by the NFV Orchestrator (NFVO). The NFVO is also capable of orchestrating end-to-end network services extending beyond a single data center.

In this new environment, the ETSI NFV has recently proposed how Service Chaining could be implemented. Service Functions can either be physical or virtual, with two options for the other required Service Chain entities proposed by ETSI NFV EVE. The first option specifies Service Chaining entities within the NFV Infrastructure (NFVI) and the second option specifies Service Chaining entities within the tenant domain (i.e. as VNFs).

Nokia Service Chaining solution target architecture

The Nokia solution for Service Chaining is pluggable on top of any Infrastructure-as-a-Service (IaaS), with no dependencies on an operator’s Evolved Packet Core (EPC) installed base. This enables operators to focus on the applications and services they are interested to integrate on top of the Service Chaining framework. Service Functions are recommended to be virtualized as well as being capable of interworking with the NSH header. In case of incompatibility with NSH, the Nokia Service Chaining framework offers a relevant proxy for interworking with the NSH header. Integration to legacy and even physical Service Functions is also possible.

The Nokia solution is compatible with all relevant Service Chaining standardization activities. The basic architecture followed is the one proposed by the IETF. With respect to the two alternatives described by ETSI NFV EVE, Nokia has chosen the second option: Service Chaining in the tenant domain with all relevant entities as guest VNFs. This option ensures the Nokia solution is cloud stack and data centre SDN vendor agnostic, otherwise tight integration would be needed for both the Virtual Infrastructure Manager (VIM) of the NFVI and the data center SDN controller in the operator’s NFVI. Regarding the enforcement point of traffic steering policies, Nokia has chosen the approach with a new reference point St and the TSSF entity. This allows the deployment of Service Chaining without affecting the existing EPC solution of the operator (apart from the need to enhance PCRF). Nokia believes the introduction of the TSSF has the following benefits:

- The controller has an end-to-end view of the Service Chaining enabled domain with respect to the instances of Service Functions/SFFs, their network topology, load/state, lifecycle operations, KPIs/attributes and lifecycle operations of the network links connecting the Service Functions. This end-to-end view enables the controller to construct Service Function
Paths in the most efficient way. A similar capability is not available when the PCEF or TDF are acting as traffic steering policy enforcement points. Both the PCEF/TDF are based on a static topology unless their architecture role changes severely so that they become controllers of the Service Chaining enabled domain.

- The controller allows dynamic discovery of Service Functions and their location.
- The controller supports dynamic instantiation of Service Function Paths (i.e. Service Function instances are selected according to their states and attributes at the time of demand, specifically at initial classification or during intermediate traversal of the Service Function Path).

The architecture and building blocks of the Nokia Service Chaining solution is illustrated in Figure 7.

The following entities are involved:

- **Service Orchestrator** - defines service steering policies and Service Chains
- **NFV Orchestration** - manages the network service’s life-cycle and orchestrates the creation of all the relevant Service Chain components (i.e. Service Chain Controller (SCC), Classifier, SFFs, Service Chain Functions (SCFs), networking) required to run Service Chains
- **Cloud Application Manager (CaM)** - manages the lifecycle operations of the various VNFs (e.g. SCFs, SCC, Layer 7 Classifier, SFF)
- **PCRF** - acts as the policy decision point of traffic steering policies (as specified in 3GPP Release 13 FMSS feature)
• SCC
  o Embraces the TSSF functionality
  o Maps a subscriber’s traffic steering policy to the Service Function Paths taking into consideration the current data center topology (with input from the NFVO)
    o Controls (optionally via embedded SDN controller) the SFFs and the Classifier
• Classifier (Class.) - performs Layer 7 analysis and classification of data flows, inserting subscriber specific metadata and Service Chain routing information into the user plane following the instructions of the SCC
• SFF - performs Layer 3 and Layer 4 classification of data flows, inserting subscriber specific metadata and Service Chain routing information into the user-plane following the instructions of the SCC. Acts as a proxy of legacy service functions to provide metadata via interfaces they support and/or handle the NSH header
• Nokia Service Chain API - enables integration of third party Service Functions to the Nokia Service Chaining solution (e.g. on NSH, utilized by the SFFs)
• SCF - a collection of virtual and/or physical appliances implementing specific user plane services

The Nokia Service Chaining solution also supports Service Chains extended beyond a single data center. The NFVO’s role is mandatory in such a use case, as the NFVO orchestrates the SDN controller of the inter-Wide Area Network (WAN) domain for inter data center connectivity and coordinates the SCC’s of each individual data center.

Conclusion

This paper has described the concepts of Service Chaining in the context of mobile networks and has outlined the solution alternatives to support Flexible Service Chaining and architecture variants. Flexible control of the Service Chains to setup a chain based on subscription information or network conditions, and to modify the chain on demand are key enablers for operators to provide Value Added Services to their customers. Flexible Service Chaining also ensures services are provided in a cost efficient manner through the efficient use of network resources that can scale on demand.

Future mobile networks are assumed to be built on cloud techniques, utilizing Virtualized Network Functions running in data centers and on SDN principles. Therefore it is a natural choice for operators to base Flexible Service Chaining solutions on these new paradigms. Flexible Service Chaining will become totally flexible when adopting an architecture that is implemented in a Telco cloud based with SDN.
The Nokia solution for Service Chaining is pluggable on top of any IaaS with no dependencies on an operator's EPC installed base. This enables operators to focus on the applications and services they are interested to integrate on top of the Service Chaining framework.

References

1. 3GPP TR 23.718: Architecture Enhancement for Flexible Mobile Service Steering
2. IETF SFC: [https://datatracker.ietf.org/wg/sfc/documents/](https://datatracker.ietf.org/wg/sfc/documents/)
3. Open Network Foundation: [https://www.opennetworking.org/](https://www.opennetworking.org/)