

Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 4, 2017

J. Dong
S. Bryant
Huawei Technologies
October 31, 2016

Problem Statement of Network Slicing in IP/MPLS Networks
draft-dong-network-slicing-problem-statement-00

Abstract

The research and standardization of IMT-2020 (a.k.a. 5G) are in progress in several industry communities and standard organizations. The goal of 5G is to integrate various services, each of which has a set of unique requirements, into a single network, such that each service has a customized network suited to its needs. The concept "Network Slicing" is widely discussed and considered as the key mechanism to meet the diverse service requirements concurrently with the same physical network infrastructure. This document provides an overview of the concept "network slicing" in the current IMT-2020 (a.k.a. 5G) related works, and discusses the corresponding requirements on IP/MPLS network, which will be used as the mobile transport network for 5G.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 4, 2017.

Internet-Draft

Network Slicing Problem Statement

October 2016

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Network Slicing Problem Statement	3
2.1.	Isolation and Separation	3
2.2.	Customization of the Topology	5
2.3.	Flexibility of the Topology	7
2.4.	Guaranteed Quality of Service	7
2.5.	Management Considerations	8
3.	IANA Considerations	8
4.	Security Considerations	8
5.	Acknowledgements	9
6.	Informative References	9
	Authors' Addresses	10

[1.](#) Introduction

The research and standardization of IMT-2020 (a.k.a. 5G) are in progress in several industry communities and standard organizations. The goal of 5G is to integrate various services, each of which has a set of unique requirements, into a single network, such that each service has a customized network suited to its needs. The concept "Network Slicing" is widely discussed and considered as the key mechanism to meet diverse service requirements concurrently with the same physical network infrastructure.

The Next Generation Mobile Networks (NGMN) gives the definition of network slice in [[Network-Slicing-Concept](#)]:

"Network Slice Instance: a set of network functions, and resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the Service Instance(s)."

[TR23.799] of 3rd Generation Partnership Project (3GPP) identifies the support of network slicing as one of the key issues to be resolved in the NextGen system:

"Network slicing enables the operator to create networks customised to provide optimized solutions for different market scenarios which demands diverse requirements, e.g. in the areas of functionality, performance and isolation."

In Focus Group (FG) IMT-2020, which is the Focus Group in ITU-T working on 5G transport network, network slicing is discussed in the Network Softwarization work item. [[FG-IMT2020-Gaps](#)] gives the definition of slicing: Slicing allows logically isolated network partitions (LINP) with a slice being considered as a unit of programmable resources such as network, computation and storage.

In order to meet the diverse service requirements, end-to-end network slicing is required in 5G, which includes the slicing of the User Equipment (UE), Radio Access Network (RAN), mobile Core network and also the mobile transport network. As one of the widely deployed mobile transport networks, IP/MPLS networks need to provide the functionality and capability required by network slicing.

[2.](#) Network Slicing Problem Statement

This section analyzes the requirements of network slicing on IP/MPLS networks, and identifies the potential gaps between the existing mechanisms and the network slicing requirements.

In IP/MPLS networks, Virtual Private Network (VPN) has been widely deployed to provide many different virtual networks on the same physical operator network. It would be beneficial to reuse the existing VPN technologies when possible, with some enhancements from the newly developed technologies such as SDN, NFV, SFC etc., to meet the network slicing requirements. However, the method used to share the resources of the underlying network with the VPNs results in

competition for resources between the VPNs, which can make it difficult to provide the degree of isolation and performance needed by some services. These issues are explored in greater detail in the following sections.

[2.1.](#) Isolation and Separation

Network slicing provides a method that allows services with diverse requirements to be provided on the same physical network with greater independence than is usually provided in a packet switching network. Each network slice appears to its users as an independent, dedicated private network which is impervious to anything that is happening on

any of the other network slices. This requires a higher degree of isolation than is found in a conventional VPN where traffic patterns in one VPN can increase the latency and jitter in another VPN, and where the shared control plane means that a high workload servicing one VPN can result in less responsiveness to another VPN. The isolation and other service requirements of each user service are likely to be different, and it is important that this is represented in the slices that carry these services to provide an efficient and economic network design.

Where 5G is used as the bearer service for real time traffic in applications such as Autonomous Driving, Virtual Reality or industrial control, there is a requirement for ultra-low transport latency and guaranteed bandwidth. In such cases, dedicated data-plane resources may be needed to guarantee the performance of the network slices carrying these services. This allows a high degree of isolation between the network slices so that the required performance is always met even when there is congestion or some other type of degradation occurs in other network slices sharing the same underlying packet network.

In addition to the data plane isolation requirement described above, we need to consider the control plane isolation requirements of the various network slices. As with the data plane isolation, the required degree of isolation in control plane will also depend on the application requirements. There are essentially three degrees of control plane isolation that need to be considered: dedicated control plane, hybrid control plane and shared control plane. A dedicated control plane can provide control plane performance guarantees, and

allows customization of the control functions, which may be required for the provisioning and optimization of some critical services. With a hybrid control plane, some of the control functions are dedicated to each network slice, while others are shared amongst a number of network slices. The hybrid approach provides a flexible way of achieving the balance between performance and efficiency. With shared control plane, the network slices use the same control plane functions and resources, regardless of whether their data plane are isolated or not. This results in competition between network slices for resources and thus less isolation. In this case, a high computation or high bandwidth event in one slice will result in less responsiveness in another slice.

It is anticipated that many third-party or vertical industrial networks will be created or migrated onto the 5G network. These third-party or industrial services will be provided with different network slices, and will typically have different requirements on the operation and management of their own network slice. For some of the services, the operation and management of the network slices can be

simply delegated to the network operator, as long as the performance requirements and relative isolation of the network slices can be guaranteed. This is much as happens with today's VPN networks. However, for some other services, it is expected that the owner of the service will require more control of the network slice, such as the placement of the network functions, the establishment and selection of the transport path, network resource allocation, etc. In order to meet these requirements, network needs to provide mechanisms to allow the third parties to configure, deploy and manage their own network slice, with minimal intervention from the network operator.

The different services will each have their own level of security requirements and will probably deploy different security mechanisms. For many applications the network slice must provide protection against interception of traffic or interruption of service, by unauthorised users. However security is always a balance between the performance, complexity and resources needed, and the economics of the service, including in the case of some Internet of Things (IoT) the energy consumption requirements. The security requirements of the service carried of the network slices may be markedly different and the design needs to accommodate this. What is of critical

importance is that each slice is impervious to an attack on any or all of the other network slices. Thus for example if there is a DDoS attack on the elements of one slice, there MUST NOT be any impact on the data plane or control plane of the other network slices.

The IP/MPLS network that is the bearer of these network slices needs to provide the mechanisms required to meet the diverse isolation requirements in data plane, control plane, network operation and security. Existing VPN technologies use a mixture of logical separation, and rely on network traffic engineering, either through metric tuning, RSVP, or Segment Routing to provide a degree of traffic isolation. However the isolation is only partial since the VPNs compete for the same resources. Thus to provide the enhanced degree of isolation needed to support more demanding service requirements, a greater degree of isolation needs to be provided by the packet network than is currently.

[2.2.](#) Customization of the Topology

In order to provide the bespoke network structure needed by each of the network service domains, it is necessary to provide each of the network slices with its own customized topology. There are a number of well known methods of providing a virtual topology that can be used to customize the topology:

- o Multi-topology Routing

- o Virtual Private Networks
- o Overlay Networks
- o Segment Routing

Multi-topology Routing (MTR) [[RFC4915](#)], [[RFC5120](#)], [[RFC7307](#)] is a way of causing the underlying routing layer to concurrently form multiple topologies over the physical network either applying a path metric to a link that is specified per topology and computing a shortest path tree that is customised to that topology. Another technique is to use a different ships in the night routing protocol such as Maximally Redundant Trees [[RFC7812](#)]. MRT relies on a common routing protocol and a common compute engine to maintain the topology. MRT has only limited application to specialist problems. In neither case is there

integration with the data-plane to maintain isolation between the slices. Furthermore it can be difficult to set up and maintain the metrics to get the degree of topology control needs by the various services. In both cases a characteristic of the user packet needs to be used to mark the packet into the correct topology.

VPNs are often used to create virtual topologies which separate and isolate the traffic of different users or services. In some VPNs [[RFC4364](#)] , [[RFC4761](#)] , [[RFC7432](#)] a common control plane is used to run the topology of the VPN and the topology of the bearer or transport network. Where a separate protocol instance is used, for example as a separate instance of BGP, the control plane of each instance is isolated, but the control traffic and the user traffic of all the instances normally share. If the control plane engages with a resource reservation protocol such as RSVP a further degree of isolation is possible, but this may not be sufficient for the most sensitive applications.

An overlay network is normally completely independent of the underlay that provides it with transport services, and normally with no coupling of the routing/signalling protocols and no way to reserve the resources in the underlying data plane, the required degree of isolation is not achieved for the most sensitive applications. Furthermore with this approach the applications have no control over the paths that their packets take across the network.

Segment routing (SR) [[I-D.ietf-spring-segment-routing](#)] is a technique that bears further consideration in this application space. With the strict source routing approach it is possible for an edge node to precisely specify the network path for its traffic. With loose source routing less control is available and it is a matter of further study whether this provides the degree of isolation needed in the network slicing environment. It is possible to have in effect a

control plane and topology per service with the SR approach. However there would need to be co-ordination between the entity creating the topologies and some entity managing the resources in the network. The use of this approach needs further study.

[2.3.](#) Flexibility of the Topology

As described by NGMN, a network slice is formed by a set of network

functions, and resources to run these network functions. With the introduction of Network Function Virtualisation (NFV) and Mobile Edge Computing (MEC), the network functions can be dynamically created at different locations, and can migrate from one place to another dynamically. The flexible and dynamic positioning of network functions in a network slice requires that the IP/MPLS networks be enhanced to have the capability of dynamically provisioning the customized network slice topology with on demand connectivity instantiated between the network functions.

The requirements is for existing topologies to be modified and new topologies added without any disruption to the other operating topologies. This will require particular attention to the impact on the data plane since reconfiguration of a topology of a network slice may lead to detectable changes, possibly transient, possibly permanent in the forwarding behaviour of other network slices.

[2.4.](#) Guaranteed Quality of Service

5G aims to provide diversified services on the same physical network. One important type of 5G service is mission critical communication. The typical use cases for this are autonomous driving, remote surgery and industrial control systems, etc, which currently require direct point to point communication, or a dedicated network over fixed infrastructure. These services have stringent requirements on latency, jitter, bandwidth, availability and reliability, etc. It is thus necessary that network slices used to carry mission critical services provide end-to-end guaranteed performance. In addition, some enhanced Mobile BroadBand (eMBB) services such as Virtual Reality (VR) which are also the target of 5G operators require transport latency to be at the millisecond (ms) level, and the bandwidth requirement will be several hundreds of Mbps.

With the exception of service carried over traffic engineered label switched paths (LSPs) using resource reservation for that LSP, existing VPN technologies share the resources of the underlying network with other VPNs results in competition for resources between them, which makes it difficult to provide the degree of performance needed by the mission critical services. Even when traffic engineering solutions are deployed, there is short term contention

for bandwidth making it difficult to achieve the very low latencies

some of these proposed new services demand.

[DETNET-WG] is working on the deterministic data paths over layer 2 and layer 3 network segments, such deterministic paths can provide the bounds on latency, loss, and packet delay variation (jitter), and high reliability that are required. Network slices of an IP/MPLS network may take advantage of the mechanisms defined in Detnet to meet the performance requirement of 5G services.

[2.5.](#) Management Considerations

As the sliced network evolves it will be necessary to provision, de-provision and modify network slices. Great care needs to be exercised in this so as to avoid disrupting other slices. This is a more difficult problem than we have historically addressed, except perhaps in the case of specialist time transfer services, because changes in topology can impact the latency of traffic running in the network. The temptation is to avoid this by freezing the paths of existing services. However the danger is that as the network ages, it will become stale with resources stranded because the running services are unable to be modified for fear of disrupting them, whilst new services cannot be provisioned because it is not possible to glean the resources they need from the fragments of discontinued services. Some form of dynamic garbage collection may therefore be needed that operates in such a way as not to introduce a transient into running network slices.

[3.](#) IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

[4.](#) Security Considerations

The security of traffic into and over a network slice needs to be addressed by the owner of the network slice, and it is expected that this would use state of the art methods. Because of the diversity of requirements these are outside the scope of this document.

The security of the slices themselves is an important consideration in the design and operation of the network slicing technology. It is important that an attack on the network slicing system is not used as a method of disrupting a targeted network slice, which may be of high value, or of a critical nature, possibly with safety of life consequences.

The nature of the vulnerability of a network slice may be more subtle than we are ordinarily concerned with. Given the delay sensitive nature of the traffic being carried over some network slices a relatively minor congestion or modulated congestion may be sufficient to cause disruption to the slice. It is therefore important to police the ingress traffic of all services, and to take precautions to protect any traffic metering technology deployed.

5. Acknowledgements

TBD

6. Informative References

[DETNET-WG]

"IETF Detnet Working Group", 2016,
<<https://datatracker.ietf.org/wg/detnet/>>.

[FG-IMT2020-Gaps]

"FG IMT-2020: Report on Standards Gap Analysis", 2015,
<<http://www.itu.int/en/ITU-T/focusgroups/imt-2020>>.

[I-D.ietf-spring-segment-routing]

Filsfils, C., Previdi, S., Decraene, B., Litkowski, S.,
and R. Shakir, "Segment Routing Architecture", [draft-ietf-spring-segment-routing-09](#) (work in progress), July 2016.

[Network-Slicing-Concept]

"Description of Network Slicing Concept", 2016,
<https://www.ngmn.org/uploads/media/160113_Network_Slicing_v1_0.pdf>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997,
<<http://www.rfc-editor.org/info/rfc2119>>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), DOI 10.17487/RFC4364, February 2006, <<http://www.rfc-editor.org/info/rfc4364>>.

[RFC4761] Kompella, K., Ed. and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", [RFC 4761](#), DOI 10.17487/RFC4761, January 2007,
<<http://www.rfc-editor.org/info/rfc4761>>.

Internet-Draft

Network Slicing Problem Statement

October 2016

- [RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", [RFC 4915](#), DOI 10.17487/RFC4915, June 2007, <<http://www.rfc-editor.org/info/rfc4915>>.
- [RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", [RFC 5120](#), DOI 10.17487/RFC5120, February 2008, <<http://www.rfc-editor.org/info/rfc5120>>.
- [RFC7307] Zhao, Q., Raza, K., Zhou, C., Fang, L., Li, L., and D. King, "LDP Extensions for Multi-Topology", [RFC 7307](#), DOI 10.17487/RFC7307, July 2014, <<http://www.rfc-editor.org/info/rfc7307>>.
- [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", [RFC 7432](#), DOI 10.17487/RFC7432, February 2015, <<http://www.rfc-editor.org/info/rfc7432>>.
- [RFC7812] Atlas, A., Bowers, C., and G. Enyedi, "An Architecture for IP/LDP Fast Reroute Using Maximally Redundant Trees (MRT-FRR)", [RFC 7812](#), DOI 10.17487/RFC7812, June 2016, <<http://www.rfc-editor.org/info/rfc7812>>.
- [TR23.799] "Study on Architecture for Next Generation System", 2012, <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3008>>.

Authors' Addresses

Jie Dong
Huawei Technologies

Email: jie.dong@huawei.com

Stewart Bryant
Huawei Technologies

Email: stewart.bryant@gmail.com

Dong & Bryant

Expires May 4, 2017

[Page 10]