Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 7, 2020

Z. Li
S. Peng
Huawei Technologies
D. Voyer
Bell Canada
C. Xie
China Telecom
P. Liu
China Mobile
Z. Qin
China Unicom
K. Ebisawa
Toyota Motor Corporation
S. Previdi
Individual
J. Guichard
Futurewei Technologies Ltd.

March 6, 2020

Problem Statement and Use Cases of Application-aware Networking (APN) draft-li-apn-problem-statement-usecases-00

Abstract

Network operators are facing the challenge of providing better network services for users. As the ever developing 5G and industrial verticals evolve, more and more services that have diverse network requirements such as ultra-low latency and high reliability are emerging, and therefore differentiated service treatment is desired by users. However, network operators are typically unaware of which applications are traversing their network infrastructure, which means that only coarse-grained services can be provided to users. As a result, network operators are only evolving their infrastructure to be large but dumb pipes without corresponding revenue increases that might be enabled by differentiated service treatment. As network technologies evolve including deployments of IPv6, SRv6, Segment Routing over MPLS dataplane, the programmability provided by IPv6 and Segment Routing can be augmented by conveying application related information into the network. Adding application knowledge to the network layer allows applications to specify finer granularity requirements to the network operator.

This document analyzes the existing problems caused by lack of application awareness, and outlines various use cases that could benefit from an Application-aware Networking (APN) architecture.

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 https://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 September 7, 2020.

Table of Contents

1. Introduction

Due to the requirement for differentiated traffic treatment driven by diverse new services, the ability to convey the characteristics of an application's traffic flow and program the network infrastructure accordingly to provide fine-grained service assurance is becoming increasingly necessary for network operators. The Application-aware Networking (APN) architecture is being defined to address the requirements and use cases described in this document. APN takes advantage of network programmability by conveying application related information in the data plane allowing applications to specify finer grained requirements to the network infrastructure.

2. Terminology

ACL: Access Control List

APN: Application-aware Networking

APN6: Application-aware Networking for IPv6/SRv6

DPI: Deep Packet Inspection

PBR: Policy Based Routing

QoE: Quality of Experience

SDN: Software Defined Networking

SLA: Service Level Agreement

MPLS: Multiprotocol Label Switching

SR: Segment Routing

SRv6: Segment Routing over IPv6 dataplane

SR-MPLS: Segment Routing over MPLS dataplane

VPN: Virtual Private Network

TE: Traffic Engineering

FRR: Fast Reroute

CAPEX: Capital expenditures

OPEX: Operating expenditures

3. Problem Statement

This section summarizes the challenges currently faced by network operators when attempting to provide fine-grained traffic operations to satisfy the various application-awareness requirements demanded by new services that require differentiated service treatment.

3.1. Large but Dumb Pipe

In today's networks, the infrastructure through which user traffic is forwarded is not able to determine information about the packet, including which application the traffic belongs to, without the introduction of middleware such as DPI, that is, the network and applications are decoupled. It is therefore difficult for network operators to provide fine-grained traffic operations for performance-demanding applications. In order to satisfy the SLA requirements network operators continue to increase the network bandwidth but only carrying very light traffic load (around 30%-40% of its capacity). This situation greatly increases the CAPEX and OPEX but only brings very little revenue from the carried services.

3.2. Network on Its Own

As the network evolves, technologies such as VPN, TE, FRR, SFC, Network Slicing, etc play important roles in satisfying service isolation, SLA guarantee, and high reliability, etc. These network technologies have themselves been evolving, introducing new features that forces the network operator to be continuously upgrading their

network infrastructure. However, none of these network technologies make the network aware of which application traffic belongs to and the fine granularity requirements of the application. Therefore, such continuous network infrastructure upgrade doesn't always enable true fine-grained traffic operation, therefore reducing the ability to bring corresponding revenue increase.

3.3. Decoupling of Network and Applications

MPLS played a very important role in helping the network enter the generation of All-IP successfully. However, MPLS alone doesn't allow a close interworking with the application layer since MPLS encapsulation is, typically, not used by the packet source.

As new services continuously evolve, more encapsulations are required, and this isolation and decoupling has further become the blockage towards the seamless convergence of the network and applications.

3.4. Challenges of Traditional Differentiated Service Provisioning

Several IETF activities have been reviewed which are primarily intended to evolve the IP architecture to support new service definitions which allow preferential or differentiated treatment to be accorded to certain types of traffic. The challenge when using traditional ways to guarantee an SLA is that the packets are not able to carry enough information for indicating applications and expressing their service/SLA requirements. The network devices mainly rely on the 5-tuple of the packets or DPI. However, there are some challenges for these traditional methods in differentiated service provisioning:

1. Five Tuples used for ACL/PBR: five tuples are widely used for ACL/PBR matching of traffic. However, these features cannot provide enough information for the fine-grained service process, and can only provide indirect application information which needs to be translated in order to indicate a specific application.

2. Deep Packet Inspection (DPI): If more information is needed, it must be extracted using DPI which can inspect deep into the packets for application specific information. However, this will introduce more CAPEX and OPEX for the network operator and impose security challenges.

3. Orchestration and SDN-based Solution: In the era of SDN, typically, an SDN controller is used to manage and operate the network infrastructure and orchestrator elements introduce application requirements so that the network is programmed

accordingly. The SDN controller can be aware of the service requirements of the applications on the network through the interface with the orchestrator, and the service requirement is used by the controller for traffic management over the network. However, this method raises the following problems:

A. The whole loop is long and time-consuming which is not suitable for fast service provisioning for critical applications;

B. Too many interfaces are involved in the loop, as shown in Figure 1, which introduce challenges of standardization and inter-operability.

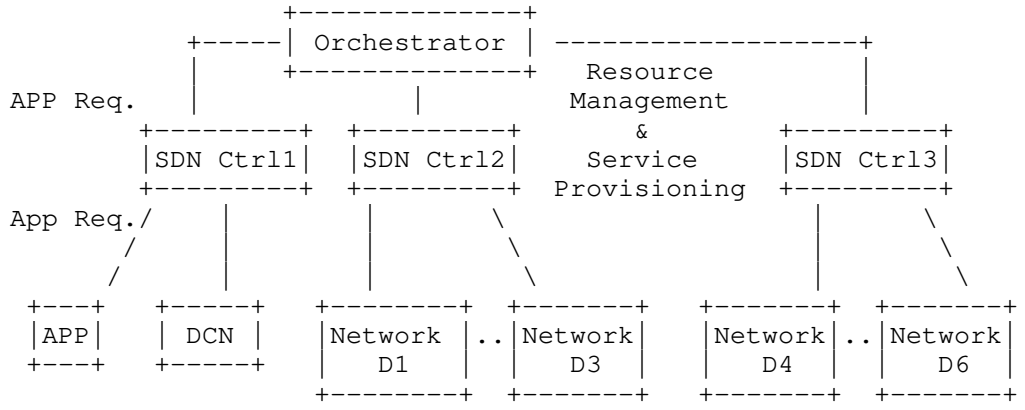

Figure 1: Multiple interfaces involved in the long service-provisioning loop

3.5. Challenges of Supporting New 5G and Edge Computing Technologies

New technologies such as 5G, IoT, and edge computing, are continuously developing leading to more and more new types of services accessing the network. Large volumes of network traffic with diverse requirements such as low latency and high reliability are therefore rapidly increasing. If traditional methods for differentiation of traffic continue to be utilized, it will cause much higher CAPEX and OPEX to satisfy the ever-developing applications' diverse requirements.

4. Key Elements of Application-aware Networking (APN)

Application-aware Networking (APN) aims to address the problems mentioned in Section 3, associated with fine-grained traffic operations that are required in order to satisfy the various

application-awareness requirements demanded by new services that need differentiated service treatment. APN aims to implement a mechanism through which application information is conveyed into the network infrastructure and that describes characteristics of the application associated with a traffic flow (e.g., application identification, network performance requirements), allowing the network to quickly adapt and perform the necessary resource adjustments so to maintain SLA performance guarantees, and hence better serve application fine-grained service requirements.

APN has the following key elements:

1. Application information is conveyed in the data plane through augmentation of existing encapsulations such as IPv6, SRv6 and MPLS. The conveyed application characteristic information (application-aware information) includes application identification and/or its network performance requirements. This element is not intended to be enforced but rather it provides an open option for applications to decide whether to input this application-aware information into their data stream. When a data packet uses APN and conveys the application information, it is referred in this document as an APN packet.

2. Application information and network service provisioning matching providing fine-granularity network service provisioning (traffic operations) and SLA guarantee based on the application-aware information carried in APN packets. This element provides the network capabilities to applications. According to the application-aware information, appropriate network services are selected, provisioned, and provided to the demanding applications to satisfy their performance requirements.

3. Network measurement of network performance and update the match between the applications and corresponding network services for better fine-granularity SLA compliance. The network measurement methods include in-band and out-of-band, passive, active, per-packet, per-flow, per node, end-to-end, etc. These methods can also be integrated.

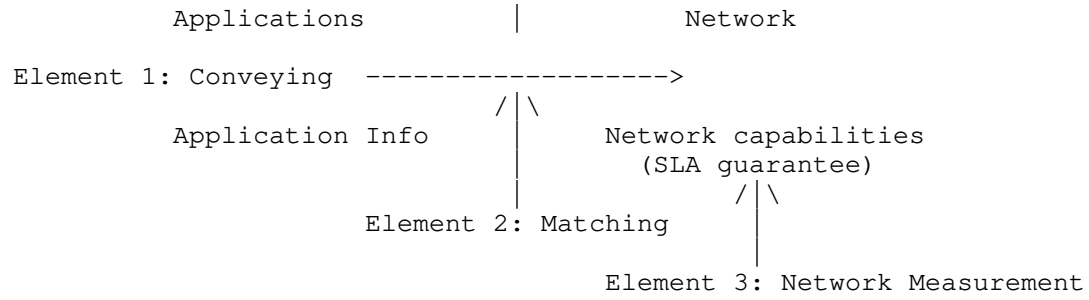

Figure 2: Illustration of the key elements of APN

4.1. Use cases for Application-aware Networking (APN)

This section provides the use cases that can benefit from the application awareness introduced by APN. The corresponding requirements for APN are also outlined.

4.1.1. Application-aware SLA Guarantee

One of the key objectives of APN is for network operators to provide fine-granularity SLA guarantees instead of coarse-grain traffic operations. This will enable them to provide differentiated services for different applications and increase revenue accordingly. Among various applications being carried and running in the network, some revenue-producing applications such as online gaming, video streaming, and enterprise video conferencing have much more demanding performance requirements such as low network latency and high bandwidth. In order to achieve better Quality of Experience (QoE) for end users and engage customers, the network needs to be able to provide fine-granularity and even application-level SLA guarantee. Differentiated service provisioning is also desired.

The APN architecture MUST address the following requirements:

- APN needs to perform the three key elements as described in Section 4.

- Support application-level fine-granularity traffic operation that may include finer QoS scheduling.

4.1.2. Application-aware network slicing

More and more applications/services with diverse requirements are being carried over and sharing the network operators' network infrastructure. However, it is still desirable to have customized network transport that can support some application's specific

requirements, taking into consideration service and resource isolation, which drives the concept of network slicing.

Network slicing provides ways to partition the network infrastructure in either the control plane or data plane into multiple network slices that are running in parallel. These network slices can serve diverse services and fulfill their various requirements at the same time. For example, the mission critical application that requires ultra-low latency and high reliability can be provisioned over a separate network slice.

The APN architecture MUST address the following requirements:

- o APN needs to perform the three key elements as described in Section 4 in the context of network slicing.
- o For the element 2, the APN architecture MUST allow to assign a given traffic flow to specific network slice according to the application information carried in the APN packet.
- o For the element 3, the APN architecture MUST allow the network measurement of each network slice.

4.1.3. Application-aware Deterministic Networking

[RFC8578] documents use cases for diverse industry applications that require deterministic flows over multi-hop paths. Deterministic flows provide guaranteed bandwidth, bounded latency, and other properties relevant to the transport of time-sensitive data, and can coexist on an IP network with best-effort traffic. It also provides for highly reliable flows through provision for redundant paths.

The APN architecture MUST address the following requirements:

- o APN needs to perform the three key elements as described in Section 4 in the context of deterministic networking.
- o For the element 2, the APN architecture MUST allow to assign a given traffic flow to a specific deterministic path according to the application information carried in the APN packet.
- o For the element 3, the APN architecture MUST allow the network measurement of each application-aware deterministic path.

4.1.4. Application-aware Service Function Chaining

End-to-end service delivery often needs to go through various service functions, including traditional network service functions such as firewalls, DPIs as well as new application-specific functions, both physical and virtual. The definition and instantiation of an ordered set of service functions and subsequent steering of the traffic through them is called Service Function Chaining (SFC) [RFC7665]. SFC is applicable to both fixed and mobile networks as well as data center networks.

Generally, in order to manipulate a specific application traffic along the SFC, a DPI needs to be deployed as the first service function of the chain to detect the application, which will impose high CAPEX and consume long processing time. For encrypted traffic, it even becomes impossible to inspect the application.

The APN architecture MUST address the following requirements:

- APN needs to perform the three key elements as described in Section 4 in the context of service function chaining.
- For the element 1, class information can be conveyed.
- For the element 2, the APN architecture MUST allow to assign a given traffic flow to a specific service function chain and MUST allow the subsequent steering according to the application information carried in the APN packets.
- For the element 3, the APN architecture MUST allow the network measurement of each application-aware service function chain.

4.1.5. Application-aware Network Measurement

Network measurement can be used for locating silent failure and predicting QoE satisfaction, which enables real-time SLA awareness/ proactive OAM. Operations, Administration, and Maintenance (OAM) refers to a toolset for fault detection and isolation, and network performance measurement. In-situ Operations, Administration, and Maintenance (IOAM) records operational and telemetry information in the packet while the packet traverses a path between two points in the network.

The APN architecture MUST address the following requirements:

- APN needs to perform the two key elements as described in Section 4 in the context of network measurement. The network measurement in the element 3 does not need to be considered here.

5. Application-aware IPv6 Networking (APN6)

As mentioned in Section 3.3, MPLS dataplane is not (or rarely) used at the packet origin (i.e., where the packet is sourced) and therefore it is not possible to assume the MPLS encapsulation is available end-to-end in the traffic flow journey. This scenario is still supported by APN with the ability to classify the packet at the ingress node of the MPLS domain. Of course, it reduces the seamless inter-working between applications and network layer but still APN will improve the resources utilization of the network layer.

APN is intended to be dataplane agnostic. Hence, APN architecture, functions and elements are applicable to both IPv6/SRv6 and MPLS dataplanes. However, it is obvious that IPv6/SRv6 dataplane delivers a better option for APN due to its flexibility, address space and later developments of SRv6 as of [I-D.ietf-6man-segment-routing-header] and [I-D.ietf-spring-srv6-network-programming]. Therefore, this document is mostly focused on the IPv6/SRv6 dataplane. MPLS dataplane is also supported by APN but with some limitations such as backward compatibility and limited address space (20 bits label size).

In this document we refer to APN6 when APN applies to the IPv6/SRv6 dataplane. Application-aware IPv6 Networking (APN6) aims to address APN problems described in Section 3 in the IPv6/SRv6 dataplane. APN6 conveys information into the network infrastructure about the characteristics of the application associated with a traffic flow (including application identification and network performance requirements), using IPv6/SRv6 encapsulation allowing the network to quickly adapt and perform the necessary network resource adjustments to maintain SLA performance guarantees, and hence better serve application fine-grained service requirements.

The advantages of using IPv6/SRv6 to support APN include,

1. Simplicity: Conveying application information with IPv6 encapsulation can just be based on IP reachability.

2. Seamless convergence: Much easier to achieve seamless convergence between applications and network since both are based on IPv6.

3. Great extensibility: IPv6 encapsulation including its extension headers can be used to carry very rich information relevant to applications.

4. Backward compatibility: On-demand network upgrade and service provisioning. If the application information is not recognized

by the node, the packet will be forwarded based on pure IPv6, which ensure backward compatibility.

5. Little dependency: Information conveying and service provisioning are only based on the forwarding plane of devices, which is different from the Orchestration and SDN-based solution which involves multiple elements and diverse interfaces.

6. Quick response: Flow-driven and direct response from devices since it is based on the forwarding plane.

6. IANA Considerations

This document does not include an IANA request.

7. Security Considerations

Since the application information is conveyed into the network, it does involve some security and privacy issues.

First, APN only provides the capability to the applications to provide their profiles and requirements to the network, but it leaves the applications to decide whether to input this information. If the applications decide not to provide any information, they will be treated in the same way as today's network and cannot get the benefits from APN.

Once the application information has been carried in the IPv6 packets and conveyed into the network, the IPv6 extension headers, AH and ESP, can be used to guarantee the authenticity of the added application information.

Any scheme involving an information exchange between layers (application and network layers in this case) will obviously require an accurate valuation of security mechanism in order to prevent any leak of critical information. Some additional considerations may be required for multi-domain use cases. For example, how to agree upon which application information/ID to use and guarantee authenticity for packets traveling through multiple domains (network operators).

8. Acknowledgements

The authors would like to acknowledge Robert Raszuk (Bloomberg LP) and Yukito Ueno (NTT Communications Corporation) for their valuable review and comments.

9. Contributors

Daniel Bernier
Bell Canada
Canada

Email: daniel.bernier@bell.ca

Liang Geng
China Mobile
China

Email: gengliang@chinamobile.com

Chang Cao
China Unicom
China

Email: caoc15@chinaunicom.cn

Chang Liu
China Unicom
China

Email: liuc131@chinaunicom.cn

Cong Li
China Telecom
China

Email: licong@chinatelecom.cn

10. References

10.1. Normative References

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

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <https://www.rfc-editor.org/info/rfc7665>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8578] Grossman, E., Ed., "Deterministic Networking Use Cases", RFC 8578, DOI 10.17487/RFC8578, May 2019, <https://www.rfc-editor.org/info/rfc8578>.

10.2. Informative References

[I-D.ietf-6man-segment-routing-header] Filsfils, C., Dukes, D., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-26 (work in progress), October 2019.

[I-D.ietf-spring-srv6-network-programming] Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-10 (work in progress), February 2020.

Authors' Addresses

Zhenbin Li
Huawei Technologies
China

Email: lizhenbin@huawei.com

Shuping Peng
Huawei Technologies
China

Email: pengshuping@huawei.com

Daniel Voyer
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Chongfeng Xie
China Telecom
China

Email: xiechf@chinatelecom.cn

Peng Liu
China Mobile
China

Email: liupengyjy@chinamobile.com

Zhuangzhuang Qin
China Unicom
China

Email: qinzhuangzhuang@chinaunicom.cn

Kentaro Ebisawa
Toyota Motor Corporation
Japan

Email: ebisawa@toyota-tokyo.tech

Stefano Previdi
Individual
Italy

Email: stefano@previdi.net

James N Guichard
Futurewei Technologies Ltd.
USA

Email: jguichar@futurewei.com