Use case and Requirements for Latency Management in Network Slices  
draft-suzuki-netslices-latency-management-00

Abstract

This draft provides a use case that addresses the need for latency  
management for end-to-end data transmissions through multiple  
domains. Specifically, examples of latency management schemes are  
described. In addition, the necessity of a common latency management  
framework and of interfaces to gather latency information between  
edges in each domain and to determine data transmission paths is  
addressed.

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 September 3, 2018.

Copyright Notice

Copyright (c) 2018 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 .............................................. 3
2. Use Case Requiring Low Latency .......................... 4
3. Problem Statement ....................................... 5
4. Types of Latency Control ................................. 6
   4.1. Centralized Management Type ....................... 6
   4.2. Distributed management type ...................... 7
5. Requirements for End-to-End Latency Management .......... 9
   5.1. Latency Management Framework ................... 9
   5.2. Interface for Gathering Latency Information ...... 9
   5.3. Interface for configuring a Data Transmission Path ... 9
6. Security Considerations .................................. 10
7. IANA Considerations ..................................... 11
8. Informative References .................................. 12
Author’s Address ........................................... 13
1. Introduction

This draft provides a use case that needs low latency and the requirements for a framework to guarantee a service-level agreement (SLA) on latency for end-to-end data transmissions in a network slice. Currently, several use cases have been discussed for network slices in the document [NetSlices-Usecase]. In addition, the architecture of network slices has been discussed in the document [NetSlices-Architecture]. In a network slice, specific resources such as a network, computing power, and storage are assigned exclusively to each user or service. Therefore, it is natural for each user of a network slice to expect an SLA for the data transmission performance.

Network slice schemes will be able to provide networks with low-latency data transmission. With regard to low latency, deterministic networks have also been discussed as a means of supporting applications requiring extremely real-time data transmissions [DetNet-Architecture]. In order to contribute to an SLA for network slice schemes, a specific use case for DetNet is described in this draft. In addition, requirements for controlling latency among multiple network domains are addressed.

In Section 2, a use case requiring end-to-end latency management is described. In Section 3, the problem of having to manage latency in order to transmit data through multiple domains is discussed. In Section 4, example methods for managing latency are shown. In Section 5, requirements that the management system should satisfy are described.
2. Use Case Requiring Low Latency

One use case requiring low-latency data transmission is the control of a vehicle from a remote control center. In recent years, autonomous cars are being actively researched and developed. Such cars will be able to drive without being controlled by a driver at almost all times. However, in some cases, these cars might stop driving. For example, when some of the sensors or cameras that detect the conditions of the car’s surroundings are damaged, it will be difficult for the car to drive autonomously. In such circumstances, the damaged vehicle could be controlled by a driver at a remote control center. The driver would have to drive the damaged vehicle using the available sensors and monitors. In order to control the vehicle safely, control data must be transmitted with low latency between the remote control center and the car.
3. Problem Statement

Currently, deterministic network architecture is being actively discussed. Specifically, schemes to control latency among multiple network domains are being considered. With regard to end-to-end latency control in network slicing, there will be two types of service. One is the selection of a specific route that is able to meet the requirements of end-to-end latency. The other is the assignment of consumable latency in each domain. In the former case, a specific route could be selected by a single network management system. In the latter case, consumable latency for each domain could be negotiated among multiple network-domain management systems. Another scheme may also be possible. In order to manage end-to-end latency through multiple network domains, a common scheme is needed.
4. Types of Latency Control

With regard to controlling end-to-end latency, there are at least two types of scheme as described above. One is based on centralized management and the other is based on distributed management. In this section, those schemes are described in detail.

4.1. Centralized Management Type

An example system composed of a sender (Tx), receiver (Rx), multiple domains, sub-management systems, and a main management system is shown in Figure 1. In this system, each sub-management system manages latency and specific routes in its respective domain. In addition, each sub-management system measures the latency for possible data transmission paths. In the figure, measured latency for each path between edges in each domain and for each connection between domains is shown. The measurement results are transmitted to the main management system. The main management system selects a specific route path between the sender and receiver to meet the latency requirement of the sender, and the selection results are transmitted to the sub-management systems. They configure a data transmission path in their domain.
4.2. Distributed management type

An example system composed of a sender (Tx), receiver (Rx), multiple domains, and sub-management systems is shown in Figure 2. In this system, each sub-management system manages latency and specific routes in its respective domain. The sub-management system that receives a latency requirement for an end-to-end data transmission requests the other sub-management systems to return possible transmission latency between edges in the domain and between domains. The sub-management system selects a specific domain path between the sender and receiver to meet the latency requirement. After that, the selection results are transmitted from the sub-management system to the other sub-management systems. Specifically, the permitted latency for each domain is transmitted to each sub-management system. The selected sub-management system determines a specific route path.
to transmit data in each domain.

Figure 2: Distributed latency management system
5. Requirements for End-to-End Latency Management

A framework for managing end-to-end data transmission latency must be designed. In addition, the interfaces to gather latency information and configure data transmission paths must be prepared to control latency. Specific requirements are briefly described below.

5.1. Latency Management Framework

As shown in Figure 1 and Figure 2, there are at least two types of latency management. One is to determine a specific data transmission path using a centralized management system. The other is to determine a specific data transmission path using a distributed management system. In order to efficiently manage latency, a common framework must be determined.

5.2. Interface for Gathering Latency Information

This interface is used to gather latency information for each domain. In the case of the centralized management system, the main management system uses the interface to gather latency information for each path between edges in each domain and for each connection between domains from all the sub-management systems. In the case of the distributed management system, one sub-management system uses the interface to gather possible transmission latency information between edges in the domain and between domains from the other sub-management systems.

5.3. Interface for configuring a Data Transmission Path

This interface is used to configure an end-to-end data transmission path. In the case of the centralized management system, the main management system uses the interface to transmit a selected data-transmission route path to the sub-management systems. In the case of the distributed management system, one sub-management system uses the interface to transmit a selected data-transmission domain path to the other sub-management systems that provided possible transmission latency information between edges in each domain and between domains.
6. Security Considerations

This document describes a use case and requirements for latency management among multiple network domains. A system to manage latency for end-to-end data transmissions could be composed of multiple sub-management systems for multiple domains. In this system, latency information between edges in each domain is gathered or exchanged to determine a data-transmission route path or domain path. It is therefore necessary to use a secure communication channel between the latency management systems.
7. IANA Considerations

This document includes no requests for IANA.
8. Informative References

[DetNet-Architecture]


[NetSlices-Architecture]

<http://tools.ietf.org/html/draft-geng-netslices-architecture-02>

[NetSlices-Usecase]

Internet-Draft    Latency Management in Network Slices        March 2018

Author’s Address

Toshiaki Suzuki
Research & Development Group, Hitachi, Ltd.
1-280 Higashi-koigakubo
Kokubunji, Tokyo 185-8601
Japan

Phone: +81-42-323-1111
Email: toshiaki.suzuki.cs@hitachi.com