srv6ops
Internet-Draft
Intended status: Standards Track
Expires: 4 September 2025

X. Gu, Ed.
X. Yi, Ed.
N. Zhang, Ed.
China Unicom

3 March 2025

Requirements and Deployments for High-Speed IoV based on SRv6

draft-gu-srv6ops-req-dep-iov-srv6-00

Abstract

This document proposes a deployment scheme for high-speed IoV by utilizing CATS, IFIT, SRv6, and network slicing technologies. Requirements and problems are discussed, and a deployment scheme is described in detail.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In high-speed Internet of Vehicles (IoV), vehicles are interconnected with road infrastructure, people, cloud services, and other traffic participants through network connections, enabling real-time data interaction and sharing. With the rapid development of big data, artificial intelligence, and communication technologies, the applications of the IoV are no longer limited to in-vehicle entertainment and navigation services. Innovative applications such as autonomous driving, remote control, and vehicle-road cooperation have emerged. These emerging business types impose more stringent and flexible differentiated requirements on network capabilities.

The application types in the IoV are mainly divided into two categories. One category consists of applications highly related to driving, such as autonomous driving, remote control, and intelligent driving services. These applications require extremely low latency guarantees to avoid traffic safety accidents. The other category includes applications unrelated to driving, such as voice communication, streaming media, and other entertainment applications, which do not demand extremely high network quality guarantees.

If the traffic of these two types of services is not differentiated in the network, it may lead to the unreasonable allocation of network resources. Delay-sensitive services may experience delays, packet loss, and other issues during data transmission, posing serious security risks. At the same time, the burst data generated by delay-sensitive services may preempt the network resources of entertainment applications, affecting the user experience. In addition, when two types of service traffic are mixed together, it is likely to cause data accumulation at network nodes, resulting in network congestion and a decline in overall network performance.

By using technologies such as Segment Routing IPv6 (SRv6), network slicing, Computing-Aware Traffic Steering (CATS), and In-situ Flow Information Telemetry (IFIT), the traffic of these two types of applications can be differentiated. On the one hand, it can meet the differentiated needs of various applications and provide rich network services for applications. On the other hand, it can improve the utilization rate of network resources and achieve the optimal resource matching.

This document aims to present a solution for high-speed vehicular network scenarios. By utilizing CATS, IFIT, SRv6, and network slicing technologies, the solution distinguishes between two types of traffic flows in vehicular networks. It provides precise service guarantees tailored to different applications, achieving optimal allocation of network resources. This approach not only enhances user experience but also ensures vehicle safety in high-speed networking environments.

2. Conventions and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here. Abbreviations and definitions used in this document:

*IoV: Internet of Vehicles.

*SRv6: Segment Routing IPv6.

*CATS: Computing-Aware Traffic Steering.

*IFIT: In-situ Flow Information Telemetry.

3. Requirements

3.1. Requirements of Applications Affecting Driving

Applications related to driving, such as autonomous driving, remote control, and intelligent driving, have extremely high requirements for network quality.

--Vehicles need to make rapid decisions in response to changes in road conditions. The network must provide extremely low latency guarantees to ensure the safety and stability of vehicle driving.

--The implementation of these applications relies on a large amount of sensor data, and the accuracy and integrity of the data directly affect the decision-making and control of the vehicles. The network needs to ensure the high reliability of data transmission to ensure that vehicles can accurately perceive the surrounding environment under various complex conditions.

--The network needs to have sufficient bandwidth to support the large amount of sensor data generated during the operation of these applications. Meanwhile, in collaborative scenarios, there is a need for big data interaction between vehicles as well as between vehicles and cloud servers. The network needs to have high bandwidth to ensure the timeliness and integrity of information interaction.

--When driving at high speeds, the position of the vehicle changes rapidly. The network should be able to switch routes quickly to ensure that the vehicle maintains a stable connection with other communication entities, so as to continuously obtain accurate information.

3.2. Requirements of Applications Unrelated to Driving

Entertainment applications such as video downloading, online music playback, and email do not affect driving and are non-delay-sensitive applications. Therefore, these applications have a relatively low priority in the IoV scenario and do not impose high requirements on network quality.

4. Problem Statement

4.1. Inability to Identify the Traffic of Different Applications

The current network’s inability to identify the traffic of different applications will lead to the unreasonable allocation of network resources, which seriously affects the overall performance of the IoV and the user experience. On the one hand, for latency-sensitive applications, low latency is a crucial factor in ensuring the system’s rapid response and accurate decision-making. Network congestion may be caused by the data generated by entertainment applications, which in turn increases the latency. As a result, vehicles are unable to make correct decisions in a timely manner, and their safety cannot be guaranteed.On the other hand, although entertainment applications have a relatively low priority, users still expect to obtain smooth audio and video services or social services. In an environment with mixed traffic, these applications may be affected due to the network resources being preempted by latency-sensitive applications, leading to problems such as audio and video stuttering and delayed information sending, which will affect the service experience.

4.2. Insufficient Coordination Between Computing and Network

Meanwhile, due to the insufficient coordination between computing and networking in the current network, when calculating the routing path, it often only focuses on factors at the network level while ignoring the status of computing resources. Or it only pays attention to the allocation of computing resources without taking the actual transmission capacity of the network into account. This one-sided routing calculation method makes the adaptation between the data transmission path and the computing resources poor, and it is unable to fully leverage the advantages of the integration of computing and networking. As a result, business processing efficiency is compromised, making it difficult to meet the increasingly diverse and demanding requirements of modern services.

5. Deployment Scheme for High-Speed IoV

A vehicular network solution based on application-network collaboration was deployed and validated for the first time in Hebei Province, China. This solution integrates several advanced technologies, including SRv6 [RFC9602] [I-D.liu-srv6ops-problem-summary], network slicing, CATS [RFC9341] [I-D.yi-cats-hierarchical-metric-distribution] and IFIT [I-D.song-opsawg-ifit-framework]. By leveraging these technologies, the solution provides a flexible resource allocation framework and ensures high-quality network service guarantees for high-speed vehicular networks by leveraging these technologies. This integrated approach optimizes resource utilization while meeting the stringent performance requirements of modern vehicular communication systems.

Figure 1 shows the architectural schematic of the deployment scheme.

```
       +-------------------------------------+
+ -----|    Computing and Network Controller  |---------+
|      +-------------------------------------+          |
|       |                                    |          |
|       |      +-------------------- +      |          |
|       |      | +-----------------+ |      |          |
+-+     |      | |Dedicated Channel| |      |          |
|V|     |      | +-----------------+ |      |          |
|e|  +-------+  | （SRv6 Policy1 with  |  +-------+  +---------+
|h|  |Network|  |   network slicing）  |  |Network|  |Computing|
|i|--|       |--|                     |--|       |--|         |
|c|  |Device |  | +-----------------+ |  |Device |  |Resource |
|l|  +-------+  | |  Shared Channel | |  +-------+  +---------+
|e|            | +-----------------+ |

+-+            | （SRv6 Policy2 with- |
               | out network slicing |
               +---------------------+
```

```
<--------------------------IFIT,CATS-------------------------->
```

Figure 1: Deployment-architecture

5.1. SRv6

By leveraging SRv6 technology, network paths can be flexibly configured and dynamically adjusted according to specific service requirements, enabling precise traffic control and customized routing.

The solution provides two types of channels for application traffic: dedicated channels and shared channels. The dedicated channel ensures high availability and can provide stable, efficient network resources for critical applications in vehicular networks. In contrast, the shared channel emulates the public channels in current networks that lack quality-of-service guarantees, which may experience congestion and packet loss during periods of high service traffic.

Through the deployment of different SRv6 policies for various types of business traffic, the network can deliver differentiated services, providing dedicated assurance for key applications and enhancing overall network performance. For mission-critical traffic related to driving, the controller enforces SRv6 Policy1, directing it to traverse the dedicated channel. For non-driving-related traffic, such as media streams or file data streams, the controller applies SRv6 Policy2, routing it through the shared channel. This approach ensures optimal resource allocation while meeting diverse service requirements.

5.2. Network Slicing

Network slicing technology enables customized configuration and management according to different application scenarios and service requirements, thereby providing differentiated quality of service. For critical applications in vehicular networks, extremely high levels of reliability and real-time performance are required. Network slicing can isolate such high-priority services from ordinary ones, creating a dedicated slice with high reliability and low latency to ensure the accurate and timely transmission of control commands, while avoiding interference from other types of traffic.

In this solution, network slicing is implemented on the dedicated channel to provide high-quality network service guarantees for driving-related applications. Based on service requirements and network conditions, slicing management policies are formulated and distributed to all nodes within the network, ensuring that the slices operate as intended. This approach optimizes resource allocation and enhances the overall performance and reliability of the network for mission-critical applications.

5.3. CATS

For latency-sensitive services, a distributed routing decision-making model is adopted, where network devices perform routing decisions. This approach enables rapid routing and fast switching, which is critical for maintaining performance in time-critical applications. For non-latency-sensitive services, a centralized routing decision-making model is utilized, with routing decisions made by a centralized controller. This model provides a global perspective, allowing for higher resource utilization across the network.

In the hybrid solution, computational and networking factors are independently distributed, avoiding large-scale modifications to network devices. The choice between centralized and distributed routing decision-making is made on an as-needed basis, providing differentiated services tailored to specific business requirements. This flexible approach ensures optimal performance and resource allocation for diverse service types.

5.4. IFIT

The IFIT technology marks feature information by inserting an IFIT header into real business packets, enabling real-time perception of network traffic, topology structures, and device states. This high-precision in-band detection capability allows the network to dynamically sense the actual requirements of various application flows, thereby providing differentiated services for different applications and achieving rational network resource allocation. Moreover, IFIT technology offers high-precision performance monitoring and fault localization capabilities for key services in vehicular networks.

In computing and network controller, relevant IFIT parameters can be configured at each compute-capable routing gateway node participating in the in-band detection process, enabling the in-band detection function. At the same time, IFIT monitoring strategies are distributed to these nodes. After a vehicle completes standard 5G registration and successfully establishes an IPv6 session, it initiates a service request. The gateway devices then report the in-band detection results, allowing performance metrics to be viewed. This approach ensures efficient and precise monitoring of network performance for vehicular services.

6. Security Considerations

TBD

7. IANA Considerations

TBD

8. References

8.1. 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>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC9602] Krishnan, S., "Segment Routing over IPv6 (SRv6) Segment Identifiers in the IPv6 Addressing Architecture", RFC 9602, DOI 10.17487/RFC9602, October 2024, <https://www.rfc-editor.org/info/rfc9602>.

[RFC9341] Fioccola, G., Ed., Cociglio, M., Mirsky, G., Mizrahi, T., and T. Zhou, "Alternate-Marking Method", RFC 9341, DOI 10.17487/RFC9341, December 2022, <https://www.rfc-editor.org/info/rfc9341>.

8.2. Informative References

[I-D.liu-srv6ops-problem-summary] Liu, Y., Voyer, D., Graf, T., Miklos, Z., Contreras, L. M., Leymann, N., Song, L., Matsushima, S., Xie, C., and X. Yi, "SRv6 Deployment and Operation Problem Summary", Work in Progress, Internet-Draft, draft-liu-srv6ops-problem-summary-04, 11 February 2025, <https://datatracker.ietf.org/doc/html/draft-liu-srv6ops-problem-summary-04>.

[I-D.yi-cats-hierarchical-metric-distribution]
Yi, X., Zhang, N., and H. Shi, "Hierarchical methods of computing metrics distribution", Work in Progress, Internet-Draft, draft-yi-cats-hierarchical-metric-distribution-01, 16 October 2024, <https://datatracker.ietf.org/doc/html/draft-yi-cats-hierarchical-metric-distribution-01>.

[I-D.song-opsawg-ifit-framework]
Song, H., Qin, F., Chen, H., Jin, J., and J. Shin, "Framework for In-situ Flow Information Telemetry", Work in Progress, Internet-Draft, draft-song-opsawg-ifit-framework-21, 23 October 2023, <https://datatracker.ietf.org/doc/html/draft-song-opsawg-ifit-framework-21>.

Authors’ Addresses

Xinrui Gu (editor)
China Unicom
Beijing
China
Email: guxr12@chinaunicom.cn

Xinxin Yi (editor)
China Unicom
Beijing
China
Email: yixx3@chinaunicom.cn

Naihan Zhang (editor)
China Unicom
Beijing
China
Email: zhangnh12@chinaunicom.cn