Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 5 May 2026

J. Mao
Z. Li
N. Geng
C. Zheng
Huawei Technologies

1 November 2025

APN Framework for Internet of Agent (IoA)
draft-mao-rtgwg-apn-framework-for-ioa-00

Abstract

With the rapid development of large model technologies in the AI field, it has become possible to develop more intelligent assistant software, which is currently referred to as AI Agents in the industry. These agents may come from different manufacturers and be deployed on different cloud platforms and regions. They need to communicate and collaborate with each other through the Internet, which is called Internet of Agents (IoA). Different interactions of AI agents have varying task requirements, which also lead to different demands on the network. This requires network providing various fine granular services for the interactions of AI agents.

This document proposes the application of the APN framework in the IoA scenario and analyzes its necessity.

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 5 May 2026.

Copyright Notice

Table of Contents

1. Introduction

With the rapid development of large language model technologies in the AI field, it has become possible to develop more intelligent assistant software, which is currently referred to as AI agents in the industry.

Various types of AI agents have emerged rapidly within a year, serving different industries and people's daily lives. Examples include deep research agents, code development agents, text generation agents, image generation agents, video generation agents, as well as agents for ordering meals, booking flights, booking hotels, and travel planning.

These agents may come from different manufacturers and be deployed on different cloud platforms and regions. They need to communicate and collaborate with each other through the internet.

For example, a user might request the marketing agent to help generate promotional materials for a new router product. The marketing agent might first ask the text generation agent to write promotional copy, then request the image generation agent to create promotional images, and subsequently ask the video generation agent to produce a product introduction video. Finally, it would collect all the promotional materials from these agents, save them on the server and notify the user to review them.

Different interactions of AI agents have varying task requirements, which also lead to different demands on the network. This requires network providing various fine granular services for the interactions of AI agents.

2. Requirements Language

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 RFC 2119 [RFC2119] RFC 8174 [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Terminology

APN: Application-aware Networking

IoA: Internet of Agents

4. The Necessity of APN Framework for IoA

4.1. What is APN

APN is a technology that enables networks to be aware of application requirements. Its framework and extension are defined in [I-D.li-rtgwg-apn-framework] and [I-D.li-rtgwg-apn-app-side-framework].

These documents propose a new framework, named Application-aware Networking (APN), where application-aware information (APN attribute) including application-aware identification (APN ID) and application-aware parameters (APN Parameters), is encapsulated at network edge devices and carried along with the encapsulation of the tunnel used by the packet when traversing the APN domain.

By APN domain we intend the operator infrastructure where APN is used from edge to edge (ingress to egress) and where the packet is encapsulated using an outer header incorporating the APN information. The APN attribute will facilitate service provisioning and provide fine-granularity services in the APN domain.

In addition, for the extension of the APN framework for application side, the APN resources of the APN domain are allocated to applications which compose and encapsulate the APN attribute in packets. When network devices in the APN domain receive packets carrying APN attribute, they can directly apply policies for these traffic flows according to the APN attribute encapsulated by applications.

APN can be applied in IoA scenarios, allowing networks to understand the requirements of each agent and its tasks on the network. This enables the network to provide services that match these requirements, such as SLA assurance, security, measurement, and visualization, thereby improving the performance, security, and maintainability of agent interactions.

4.2. IoA needs Differentiated SLA Guarantee from network

Different tasks of agent interaction have varying requirements for network transmission.

For example:

* Tool invocation tasks emphasize real-time interaction and are highly sensitive to latency and jitter, but have low bandwidth requirements.

* Text generation tasks are relatively sensitive to latency but have low bandwidth requirements.

* Image generation tasks have low latency requirements but relatively high bandwidth demands.

* Video generation tasks have low requirements for latency and jitter but very high bandwidth demands.

Each of these tasks can be further divided into real-time interaction tasks and background interaction tasks:

* Real-time interaction is used in scenarios where users actively operate the system. Since the final response needs to be presented to the user (who waits for the result in front of the device), it has higher requirements for real-time performance (i.e., latency and jitter).

* Background interaction is initiated by the agent software in the application's backend system, thus having lower requirements for real-time performance.

These different QoS requirements can be identified in traffic packets using APN. Currently, APN defines four APN parameters:

* Bandwidth Requirement Parameter

* Delay Requirement Parameter

* Delay Variation Parameter

* Packet Loss Rate Parameter

New APN parameters may need to be defined to carry some information about AI models, tools, tasks, data, or tokens for some specific IoA scenarios.

By carrying these APN parameters in the messages, the network can provide differentiated and targeted services for different intelligent agent tasks to ensure a good experience for agent interactions. For example, the network can offer fine-grained SLA guarantees, fine-grained network slicing, and fine-grained deterministic networking.

4.3. IoA needs Secure Transmission from network

Agents interact with each other through the Internet, which may require the network to provide security services.

For example, a food ordering agent on a mobile phone communicates with a payment agent on the cloud, and the payment agent communicates with a bank agent, which may be on-premises. These communications involve your personal account information and financial data, so they require a higher level of security.

The APN ID or APN parameter can be used to carry encrypted transmission requirements. The network can then create IPsec tunnels or quantum transmission channels and direct the traffic of agent interactions to these secure channels to prevent eavesdropping, tampering, and forgery.

For enterprises, they may want to enforce access control to agent interactions. For instance, some agents that can access sensitive data are not allowed to collaborate with external agents to prevent data leakage. However, some general-purpose agents are allowed to interact with external agents to improve team efficiency.

The APN ID or APN parameter can be used to carry the identifier of the security group to which an agent belongs. The network can then perform access control based on these identifiers to avoid unexpected agent interactions and information security risks.

4.4. IoA needs Measurement and Visualization from network

When intelligent agents interact through the internet, the quality of communication may need to be measured for purposes such as fault localization, troubleshooting, and visualization.

APN IDs can be used to identify AI agents, and network devices can be configured to enable network measurement functions (e.g., IFIT) for specific agents, thereby enabling in-band measurement of the interaction traffic among these agents. The network controller can visually present the communication quality of these agent interactions to the network administrators.

5. APN framework for IoA

5.1. Using APN framework in IoA scenarios

The APN architecture can be applied to IoA scenarios, as illustrated in the following diagram. The IoA service requirement that needs to be carried in the APN ID or APN parameter essentially belongs to a type of application-aware information (i.e., APN attribute).

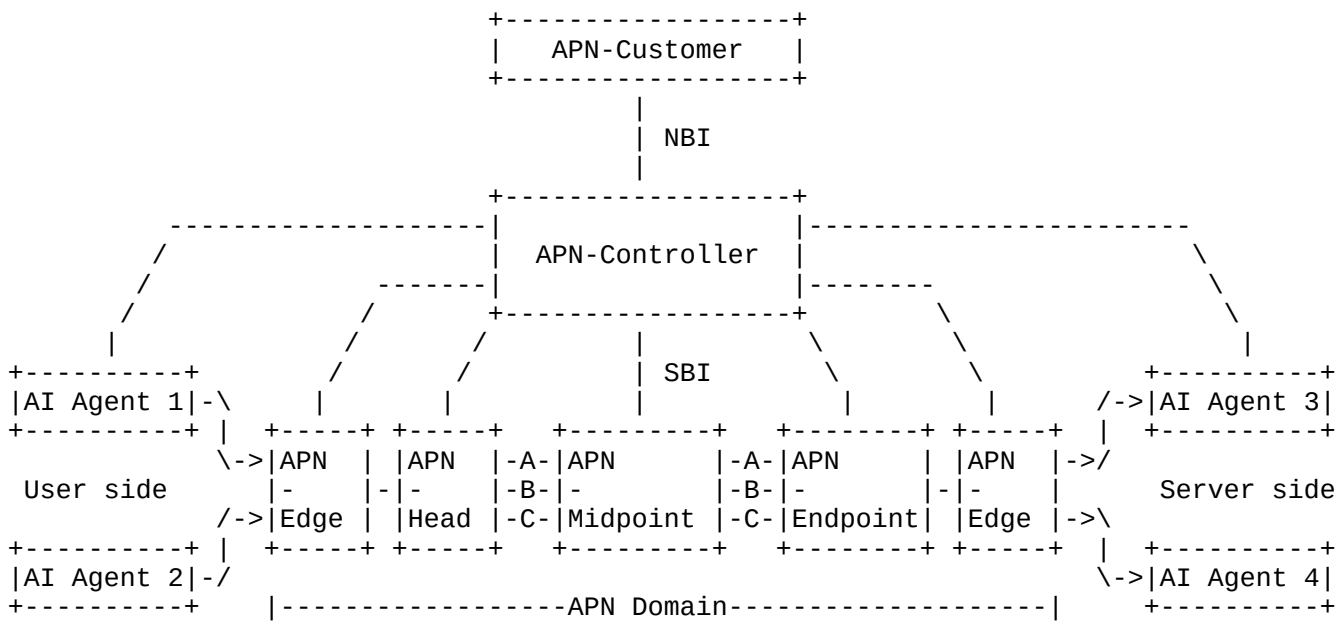

Figure 1: Using APN framework in IoA scenarios

* AI Agent: Interacts with other AI Agents on the internet, sending task requests, asynchronous notifications, and result responses to the others. It can subscribe to network services from operators and optionally carry service requirement in the sent traffic via APN Attribute. This means that the AI Agent behaves as an APN-capable Application Client or Server as defined in [I-D.li-rtgwg-apn-app-side-framework].

* APN-Edge: If AI Agents send traffic without APN attribute, APN-Edge device should obtain the APN attribute based on the configuration on this device according to the existing information in the packet header, create a tunnel encapsulation and add the APN attribute into it, and send the packets carrying the APN attribute to the APN-Head device finally. When the packets are going to leave the APN domain, APN-Edge should remove the encapsulation if APN-Edge and APN-Endpoint are deployed separately. Those are detailly defined in [I-D.li-rtgwg-apn-framework]. If the APN attribute is added by the AI Agent, the functionalities of the APN-Edge need to be changed. APN-Edge can directly transmit the packets to APN-Head without encapsulating tunnels. That is detailly defined in [I-D.li-rtgwg-apn-app-side-framework].

* APN-Head: Identifies the service requirements of IoA based on the APN attribute in the packet and provides network services that meet these requirements. This may involve encapsulating the packet with a new TE tunnel, placing the packet into an appropriate QoS queue, directing the packet to a network slice, adding or modifying some fields on the packet to perform network measurement, making access policy decisions, performing secondary encryption on the packet content, or diverting it to a secure channel.

* APN-Midpoint: APN-Midpoint provides the path service and enforces various policies according to the APN attribute carried in the packets. The APN-Midpoint may also adjust the resource locally to guarantee the service requirements depending on a specific policy and the APN attribute conveyed by the packet.

* APN-Endpoint: the process of the specific network service will end at the APN-Endpoint. If the outer tunnel header for the service between the APN-Head and the APN-Endpoint exists, it will be removed by the APN-Endpoint. If the APN attribute is sent by AI Agents, the APN-Endpoint MUST keep the APN attribute in packets without any change.

5.2. Requirements

5.2.1. Requirements for AI Agents

* [REQ A01] AI Agent should support to subscribe network services from ISP, and get service requirement identifier or the usage authorization of service requirement metadata from ISP (out-band) or its APN-Controller.

* [REQ A02] AI Agent should support to describe the service status, the quality of users’ experience (QoE), or the key quality indicators (KQI) in the APN attribute.

* [REQ A03] AI Agent should support to carry the APN attribute in the packets of the agent interaction traffic.

* TBD

5.2.2. Requirements for Network Devices

* [REQ D01] APN-Head should support steering the packets based on APN attributes into a TE tunnel that satisfies the SLA requirements of the agent interaction.

* [REQ D02] APN-Head, APN-Midpoint and APN-Endpoint should support putting the packets based on APN attributes into a QoS queue buffer that satisfies the SLA requirement of the agent interaction.

* [REQ D03] APN-Head, APN-Midpoint and APN-Endpoint should support to steer the packets based on APN attributes into a network slice that satisfies the SLA requirement of the agent interaction.

* [REQ D04] APN-Head, APN-Midpoint and APN-Endpoint should support adding, modifying, or marking some fields of the packets based on APN attributes to perform network measurement that satisfies the measurement and visualization requirement of the agent interaction.

* [REQ D05] APN-Head should support enforcing access control to the packets based on APN attributes, in accordance with some security rules or policies that satisfy the security requirement of the agent interaction.

* TBD

5.2.3. Requirements for Network Controller

* [REQ C01] APN-Controller should support to notify the allocated results of the service requirement identifier, and the usage authorization of the service requirement metadata to AI Agent.

* [REQ C02] APN-Controller should support to analysis the measure data or path data of the agent interaction traffic reported from network devices to achieve traffic visualization and network troubleshooting.

* TBD

5.2.4. Requirements for Protocol Extensions

* [REQ P01] Protocol extensions should be defined for the AI Agent to request the APN-Controller in the network to allocated the APN resources of the APN domain.

* [REQ P02] Security mechanism should be defined to guarantee for that the APN attribute being securely transmitted among the AI Agent(user side), AI Agent (server side) and the APN domain.

* TBD

5.2.5. Requirements for Security

* [REQ S01] The task contents of AI agent interactions should be protected or keep invisible.

* [REQ S02] The traffic of sensitive tasks should be steered into a secure channel for transmission, such as IPsec tunnels and quantum channels.

* TBD

6. IANA Considerations

This document does not include an IANA request.

7. Security 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>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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>.

8.2. Informative References

[I-D.li-rtgwg-apn-app-side-framework]
Li, Z. and S. Peng, "Extension of Application-aware Networking (APN) Framework for Application Side", Work in Progress, Internet-Draft, draft-li-rtgwg-apn-app-side-framework-00, 22 October 2023, <https://datatracker.ietf.org/doc/html/draft-li-rtgwg-apn-app-side-framework-00>.

[I-D.li-rtgwg-apn-framework]
Li, Z., Peng, S., Voyer, D., Li, C., Liu, P., Cao, C., and G. S. Mishra, "Application-aware Networking (APN) Framework", Work in Progress, Internet-Draft, draft-li-rtgwg-apn-framework-00, 4 March 2024, <https://datatracker.ietf.org/doc/html/draft-li-rtgwg-apn-framework-00>.

Authors' Addresses

Jianwei Mao
Huawei Technologies
Beijing
100095
China
Email: MaoJianwei@huawei.com

Zhenbin Li
Huawei Technologies
Email: robinli314@163.com

Nan Geng
Huawei Technologies
Email: gengnan@huawei.com

Chong Zheng
Huawei Technologies
Email: zhengchong6@huawei.com