RTGWG
Internet-Draft
Intended status: Standards Track
Expires: 16 September 2026

X. Zhang
China Mobile

15 March 2026

Use Cases and Requirements for AI Agent Policy-Aware Network
draft-zhang-rtgwg-agent-policy-aware-network-00

Abstract

With the widespread adoption of AI Agents, traditional network architectures can no longer meet the demand for efficient collaboration between Agents and networks. This document proposes a new paradigm of "AI Agent Policy-Aware Network", enabling three key transformations: from Flow-aware to Agent-aware, from QoS-based to Policy-intent-based, and from Network-driven to Agent-network collaborative. By defining core components such as the Agent Policy-aware Controller and Agent Policy-Aware Device, this paradigm establishes a dynamic mapping mechanism between Agent intents and network policies, supporting key scenarios including autonomous performance measurement, path optimization, SLA assurance, and secure transmission. This document outlines the background, scenarios, use cases and requirements of Agent Policy-aware Network.

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 16 September 2026.

Copyright Notice

Table of Contents

1. Introduction

Service Agent (SA): An autonomous intelligent entity capable of perceiving environmental information, making independent decisions, and executing tasks to achieve specific goals, which can interact with networks to express policy intents and obtain network resources and service guarantees.

AI Agent Policy-Aware Network: A network paradigm that can identify, understand, and execute the policy intents of AI Agents, realizing dynamic mapping between Agent intents and network policies, and supporting bidirectional collaborative decision-making between Agents and networks.

Agent Policy-Aware Controller (APC): A core component responsible for parsing Agent policy intents, dynamically mapping them to network policies, and coordinating resource allocation and policy enforcement across the network.

Agent Policy-Aware Device (APAD): A network device that can receive and execute policies issued by the Agent Policy-aware Controller or Service Agent, supporting real-time policy adjustment and enforcement based on network status and Agent requirements.

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

1.2. Terminology

Service Agent (SA): An autonomous intelligent entity capable of perceiving environmental information, making independent decisions, and executing tasks to achieve specific goals, which can interact with networks to express policy intents and obtain network resources and service guarantees.

AI Agent Policy-Aware Network: A network paradigm that can identify, understand, and execute the policy intents of AI Agents, realizing dynamic mapping between Agent intents and network policies, and supporting bidirectional collaborative decision-making between Agents and networks.

Agent Policy-Aware Controller (APAC): A core component responsible for parsing Agent policy intents, dynamically mapping them to network policies, and coordinating resource allocation and policy enforcement across the network.

Agent Policy-Aware Device (APAD): A network device that can receive and execute policies issued by the Agent Policy-aware Controller or Service Agent, supporting real-time policy adjustment and enforcement based on network status and Agent requirements.

2. Scenarios of Agent Policy-aware Networks

2.1. Medical Emergency Rescue

When a home health agent detects an elderly person fainting, it needs to immediately initiate a medical rescue process. During the rescue, vital sign data (such as heart rate and blood pressure) must be transmitted to the hospital in real time through an exclusive low-latency channel to ensure the timeliness of medical decisions. At the same time, the channel must meet high security level requirements to prevent data leakage and strictly prohibit cross-domain

transmission to comply with medical privacy regulations. The network system should dynamically create and maintain this channel without manual intervention to support efficient response within the "golden rescue time" and achieve seamless collaboration between the Agent and the network.

2.2. Game Multi-Agent Collaboration

In real-time game battles, player Agents need to obtain low-latency and low-jitter network guarantees during highly dynamic operations such as team battles and skill casting. Agents should be able to independently negotiate path switching with network Agents, supporting "lossless switching" requirements to avoid lag, while dynamically adapting to changes in game load. Based on Agent policy intents (e.g., "skill casting"), the network needs to real-time optimize transmission paths to ensure smooth operation responses. The entire process requires no human intervention, reflecting the collaborative capability of Agent active decision-making and network dynamic response, and improving the real-time and immersion of the game experience.

2.3. Government Cross-Domain Secure Transmission

Government departments need to securely transmit sensitive medical data (such as cross-provincial health records), requiring the network to automatically establish an ultra-high-security channel. The security level of the channel must be dynamically adapted based on the trust domain to which the Agent belongs (e.g., enabling group security mechanisms within the health system), avoiding manual configuration of whitelists or security policies. The network should be able to identify the identities of collaborating parties (such as digital agents of provincial health commissions and municipal medical insurance bureaus), and enable preconfigured security policies to ensure compliant and efficient data transmission. At the same time, the channel needs to support cross-domain collaboration but be strictly limited to authorized trust scopes to meet the rigid requirements of government data security and compliance.

3. Use cases

This section illustrates some use cases for Agent Policy-aware Networks.

3.1. Agent Policy-aware Performance Measurement

Service Agents have strict latency and jitter requirements for specific data flows (e.g., emergency rescue video streams), and need real-time end-to-end network quality perception to ensure business SLA compliance.

Agent Actions:

1.Collaborate with Measurement Agents on APAD, and issue structured measurement intents based on business SLA requirements (latency, jitter, packet loss, reliability).

2.Receive structured measurement results fed back by APAD upon task initialization, service handover or network quality degradation, and extract quantitative performance characteristics of business flows (total session traffic, packet length, traffic time-series patterns, etc.).

3.Predict traffic and security requirements based on measured performance characteristics, map business semantic actions to network traffic features, and output explicit QoS requirements (low latency, anti-jitter, bandwidth reservation, ultra-high security, etc.) to the network.

Network Actions:

1.Measurement Agents on APADs automatically select adaptive measurement protocols (e.g., in-band flow detection, TWAMP) according to Agent measurement intents.

2.APADs structure and feed back collected performance data to Service Agents in a standardized format.

3.Generate unique Agent policy-aware identifiers for business flows after receiving QoS requirements, and allocate dedicated network resources to implement differentiated quality guarantees.

3.2. Agent Policy-aware Path Optimization

When congestion or faults occur on the current transmission path, service SLAs may be degraded. The network should switch to a more optimal path without service interruption, enabling fault self-healing and experience self-optimization, thereby ensuring continuous service delivery.

Agent Actions:

Service Agents receive network feedback of quality degradation alarms, and issue a request for "path recalculation and switching" to the network in combination with service tolerance (e.g., "temporary interruption allowed" or "lossless switching mandatory").

Network Actions:

The Agent Policy-aware Controller (APAC) calculates an optimal set of schedulable paths that meet constraints based on real-time network topology and link status, confirms the switching strategy with Service Agents, and finally executes seamless path migration. Schedulable Path Set Information Model: -Basic Identifier: Path ID / Path Group ID -Performance Attributes: Latency, jitter, packet loss, bandwidth, remaining bandwidth -O&M Attributes: Cost, priority, reliability, congestion probability -Constraint Attributes: Slicing support, cross-domain support, encryption support

3.3. Agent Policy-aware SLA Level Assurance

Different types of traffic from Agents feature high dynamism, burstiness and periodicity (e.g., scheduled big data synchronization). The network needs to allocate exclusive resources such as bandwidth on demand to realize "tidal scheduling" of network resources, greatly improve resource utilization, and ensure deterministic experience for critical services.

Agent Actions:

Service Agents predict future traffic trends and send "resource reservation intents" (including time window, bandwidth peak, and duration) to the network.

Network Actions:

The Resource Agent of APAC evaluates the global network load: if resources are sufficient, it temporarily locks the corresponding bandwidth slice; if resources are tight, it negotiates a degradation scheme with the Agent or suggests adjusting the time window.

The Path Agent of APAC completes the full lifecycle management of bearer tunnels such as SRv6 Policy and network slicing, adjusting the level of SLA assurance and dismantling tunnels on demand.

For Agent collaboration requirements that transmit information in a specified sequence, SRv6 service function chain evolution technology is adopted to ensure the order and reliability of information transmission, meeting the timing requirements of collaborative tasks.

3.4. Agent Policy-aware Security

When Agents in different trust domains conduct data interaction, the network needs to dynamically match the corresponding security levels and encryption mechanisms to realize identity and context-based dynamic zero-trust security, ensuring that data only flows within authorized scopes.

Agent Actions:

The communicating Agents exchange identity credentials and trust domain attributes, and declare the required security levels (e.g., "national cryptographic algorithm encryption", "cross-domain isolation").

Network Actions:

The Network Security Agent parses the identities of both parties, automatically retrieves and matches preconfigured cross-domain mutual trust policies, and provides hierarchical security technologies on demand, including IPSec, QKD key distribution, and group secure transmission. The policies are automatically revoked after the session ends.

4. IANA Considerations

TBD.

5. Security Considerations

TBD.

6. References

6.1. 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/rfc/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/rfc/rfc8174>.

6.2. Informative References

[I-D.yang-rtgwg-arn-framework-04]
Yang, F. and C. Lin, "Application-Responsive Network Framework", Work in Progress, Internet-Draft, draft-yang-rtgwg-arn-framework-04, 29 June 2025, <https://datatracker.ietf.org/doc/html/draft-yang-rtgwg-arn-framework-04>.

[I-D.li-rtgwg-apn-framework]
Li, Z., Voyer, D., Li, C., Liu, P., Cao, C., Mishra, G. S., and N. Geng, "Application-aware Networking (APN) Framework", Work in Progress, Internet-Draft, draft-li-rtgwg-apn-framework-01, 12 November 2025, <https://datatracker.ietf.org/doc/html/draft-li-rtgwg-apn-framework-01>.

Author’s Address

Xiaoqiu Zhang
China Mobile
China
Email: zhangxiaoqiu@chinamobile.com