Motivations and Problem Statement of Agentic AI for network management

draft-hong-nmrg-agenticai-ps-00

Y-G. Hong (Daejeon Univ.), J-S. Youn (DONG-EUI Univ.)

Q. Wu(Huawei), B. Claise(Everything OPS)

nmrg Meeting@IETF 124 – Montreal
November 4. 2025

History and status

- 1st presentation : Motivation and problem statement of Agentic AI (NMRG 83rd Online, Oct. 8. 2025)
- 00 version I-D : draft-hong-nmrg-agenticai-ps-00 (Oct. 2025)
- 2nd presentation : Motivation and problem statement of Agentic AI (IETF 124, Nov. 2025)

Object of this draft

- 1. Outlines the key objectives of introducing Agentic AI to the field of network management
- 2. Highlight the fundamental issues with existing technologies that must be addressed to achieve these goals
- 3. Emphasizes the necessity for relevant groups within the IETF/IRTF and presents the core technological areas requiring standardization
- 4. The aim of Agentic AI is to facilitate a paradigm shift in which multiple autonomous AI agents collaborate to fully automate network operation, management and security.

Agentic AI in network management

- Paradigm shift of network management
  - Unlike traditional automation, which relies on pre-programmed rules, agentic AI applies **<u>autonomous decision-making capabilities</u>** to swiftly and accurately handle large-scale network activities and customer requests.
  - These agents continuously perform tasks such as network traffic management, fault resolution, and customer interaction support, executing responses that previously required manual human review or escalation.
- Agentic AI for network management
  - Agentic AI utilizes **<u>large language models (LLMs)</u>** to encompass a broad range of capabilities beyond natural language processing, including reasoning, problem-solving, interaction with external environments, and performing actions.
  - It possesses **<u>the capability for task decomposition</u>**, breaking down complex objectives into specific tasks and subtasks, and **<u>this cognitive capacity</u>** enables a persistent cognitive cycle (observation, inference, action) that continuously aligns network operations with high-level business intent.

Operation of Agentic AI for Network Management

```
+----------------------------------------------------------------+
|                                                                |
|   +--------------------------+                                 |
|   |  1.GOAL / INTENT(Input)  |                                 |
|   +------------------|-------+                                 |
|                      v                                         |
|            +---------+----------------+                        |
|            |    2.AI AGENT(Brain)     |                        |
|            |  (LLM/Reasoning Engine)  |                        |
|            +---|-----------------|----+                        |
|                |  Memory/Context |    |                        |
|                +-----------------+    |                        |
|                ^                      v                        |
|      +---------|----------+    +------|------------+           |
|      | 4.REFLECT(Compare) |    |  3.PLAN(Sequence) |           |
|      | (Evaluate Outcome) |<---| (Action Breakdown)|           |
|      +---------|----------+    +------|------------+           |
|                |                      v                        |
|                |   +------------------+----------------------+ |
|                +---| 5.EXECUTE(Action) via Tool Orchestrator | |
|                    +------------------|----------------------+ |
|                                       v                        |
|                 +---------------------+-----------------+      |
|                 | 6.TOOL USE(API Calls & Configuration) |      |
|                 |     (RESTCONF, Monitoring, etc.)      |      |
|                 +-----------|-------------------|-------+      |
|                             v                   ^              |
|                 +-----------+-------------------+--------+     |
|                 |   7.NETWORK ENVIRONMENT (The World)    |     |
|                 | (Apply Changes & Sense/Observe State)  |     |
|                 +----------------------------------------+     |
|                                                                |
+----------------------------------------------------------------+
```

Problem Statement of existing Techniques for Agentic AI (1/6)

- P1: Architectural bottlenecks and the failure of centralization
  - Existing AI and automation systems have often relied on centralized infrastructure for data aggregation and heavy computing. However, these centralized models **cannot handle the volume, velocity, and distributed nature of Agentic AI workloads**.
  - Centralized AI systems are constrained by central infrastructure, resulting in high latency due to round-trip times to the cloud. Such latency is unacceptable for real-time applications such as self-healing and 5G slicing management. There is also the issue that the central server becomes a bottleneck, limiting scalability.
  - The inherent limitations of such centralized models (single point of failure (SPoF), latency) inevitably **drive Agentic AI architectures towards a distributed mesh form**.
  - This leverages local processing at the edge for low latency and fault tolerance, requiring the standardization of distributed control and communication mechanisms that transcend conventional centralized SDN/management models.

Problem Statement of existing Techniques for Agentic AI (2/6)

- P2: Absence of agent-to-agent (A2A) semantic interoperability
  - Agentic systems are often built by different vendors using various frameworks, leading to fragmented and siloed system operations.
  - Complex network management tasks require the decomposition of work and collaboration between specialized agents. Without standardized agent-to-agent (A2A) protocols, bespoke connectors become necessary to connect these fragmented systems, slowing down development and integration speeds.
  - Standardization must **define consistent payloads and interfaces** that support real-time interactions between systems, **enabling agents to discover, understand, and collaborate with one another** regardless of their underlying implementations.

Problem Statement of existing Techniques for Agentic AI (3/6)

- P3: Lack of dynamic trust and accountability in autonomous behavior
  - The introduction of AI agents as autonomous entities performing actions at machine speed presents significant security and governance challenges.
  - Traditional identity and access management (IAM) focuses on human users or predefined roles. However, autonomous agents **<u>operate with dynamic intent, require context-aware access, and must maintain provable accountability for every action</u>** they perform.
  - Without a robust Zero Trust framework specifically designed for non-human autonomous entities, there is a risk of catastrophic security breaches or manipulation where autonomous systems could outpace human control capabilities.

Problem Statement of existing Techniques for Agentic AI (4/6)

- P4: Real-time data validity and resilience issues
  - The decision-making of AI agents is determined by the quality of the data they receive. In a network environment, data quality is of paramount importance.
  - Incomplete, delayed, or corrupted data feeds can lead to severe operational or financial losses when agents take autonomous actions (e.g., traffic rerouting, forced execution of financial transactions).
  - Therefore, it must extend beyond the current focus on bandwidth and speed to **include quality verification of the data agents rely upon and resilience of service paths**. This is essential to meet the requirements of continuously operating intelligent agents.

Problem Statement of existing Techniques for Agentic AI (5/6)

- P5: Problems with the existing IBN system: Rigidity of the Intent Translation Engine (ITE)
  - Existing IBN systems rely on the Intent Translation Engine (ITE) or the Intent-Based System (IBS) spatial functionality to bridge the gap between the intent user and the network operational infrastructure. This translation is typically driven by predefined data models such as YANG models and lacks the necessary adaptive flexibility when unforeseen conditions arise.
  - IBN fundamentally shifts operational modes to a dynamic intent-based approach, **yet retains the inherent limitation that control remains under human oversight**.
  - Agentic AI **minimises or eliminates human intervention** in this cognitive loop through LLM-based reasoning and planning capabilities, refining the IBN closed loop by integrating continuous reasoning and conflict resolution capabilities into the cognitive layer.
  - These capabilities represent what was lacking in the classical IBN definition and form the core technical objective.

Problem Statement of existing Techniques for Agentic AI (6/6)

- P6: ANIMA ASA(Autonomic Service Agent) 's Problem: Cognitive Simplicity
  - ANIMA's ASAs are typically designed for specific, localized autonomous functions (e.g., prefix management, bootstrapping). They rely heavily on predefined policy structures and **lack the complex reasoning, planning, or self-reflection capabilities characteristic of Agentic AI** (LLM-based task decomposition).
  - ANIMA's ASA is conceptually a precursor to Agentic AI, but lacks a cognitive core (LLM/inference engine).
  - Agentic AI introduces LLM-based planning and tool-use capabilities that require complex, semantic negotiation (A2A) beyond simple information exchange (GeneRic Autonomic Signaling Protocol; GRASP), demonstrating the necessity for a dedicated protocol layer that extends beyond the existing ANIMA framework.

Objectives of Agentic AI for Operations & Management (1/5)

- Objective 1 - Autonomous Network Operations & Management
  - Beyond minimizing human intervention, it must implement a **“Self-Driving” network that autonomously recognises, diagnoses, infers, and resolves issues** even in unpredictable situations.
  - Key Features:
    - **Predictive & Proactive Fault Management**: AI agents learn traffic patterns, logs, and performance metrics in real time to identify potential causes before faults occur. The network autonomously reroutes traffic or reallocates resources to prevent service interruptions at source.
    - **Intelligent Root Cause Analysis**: In complex, intertwined fault scenarios, multiple agents collaborate to synthesize distributed data. They deduce the root cause as a “problem of correlations” rather than a single point of failure and propose solutions.
    - **Autonomous Configuration & Optimization**: AI agents comprehend high-level objectives such as ‘optimize user experience’ and autonomously configure and continuously fine-tune routing protocols, QoS policies, security rules, and other elements to achieve them.

Objectives of Agentic AI for Operations & Management (2/5)

- Objective 2 - Intelligent & Dynamic Resource Orchestration
  - To address unpredictable traffic demands such as 6G, holographic communications, and large-scale IoT, **network resources (computing, storage, bandwidth) are allocated and coordinated in real time and proactively**.
  - Key Features:
    - **Dynamic Network Slicing**: AI agents recognize application requirements (latency, bandwidth, etc.) in real time, instantly creating, scaling, and downsizing customized network slices per user or service.
    - **Cross-Domain Resource Negotiation**: AI agents distributed across networks of different telecommunications or cloud providers negotiate in real time to dynamically secure optimal resources, ensuring end-to-end quality for global services.
    - **Edge Computing Resource Optimization**: By predicting edge node load and user mobility, AI agents dynamically reallocate workloads to optimal edge nodes while ensuring service continuity.

Objectives of Agentic AI for Operations & Management (3/5)

- Objective 3 - Predictive & Adaptive Network Security
  - Beyond defending against known attack patterns, AI agents **autonomously detect unknown zero-day attacks or advanced persistent threats (APTs) and reconfigure defence systems** in real time.
  - Key Features:
    - **Autonomous Threat Hunting & Response**: Security agents continuously detect minute anomalies across the entire network. Should an anomaly be deemed a threat, they perform immediate responses -such as isolating infected nodes or blocking attack traffic - without human intervention.
    - **Dynamic Defense Posture**: AI agents dynamically modify firewall policies, access control lists (ACLs), and traffic filtering rules in real time based on attack type and intensity, thereby minimizing the attack surface.

Objectives of Agentic AI for Operations & Management (4/5)

- Objective 4 - Enabling Novel Network Service Models
  - By transforming the network itself into **a single, vast distributed AI platform**, it enables new communication services and business models that were previously impossible.
  - Key Features:
    - **Intent-driven Service Creation**: When a user requests in natural language, ‘I want to play a lag-free VR game with my friends,’ an AI agent interprets this and provides a Network-as-a-Service that instantly allocates the necessary resources (such as network slices and edge servers).
    - **Semantic Communication**: Communication focuses on the “meaning” or “purpose” conveyed by data rather than the bits themselves, enabling ultra-efficient communication that achieves maximum effect with minimal data transmission.

Objectives of Agentic AI for Operations & Management (5/5)

- Objective 5 - Autonomous, High-Fidelity & Action-Aware Network Measurement
  - To turn raw network telemetry into trustworthy, context-rich insight that continuously retrains itself, **explains its own uncertainty, and feeds closed-loop control without human analysts**.
  - Key Features:
    - **Semantic Anomaly Narratives & Root-Cause Metrics**: Instead of threshold alerts, the model outputs human-readable stories. Each sentence is back-traced to verifiable measurement samples.
    - **Self-Driving Measurement Campaigns**: The AI translates high-level intents ("tell me if user-perceived 4 K latency could exceed 150 ms during the next football final") into dynamic sampler schedules, probe paths, and packet structures; it launches the campaign, stops when statistical confidence is reached, and releases resources back to the data plane.
    - **Counterfactual & Predictive "What-if" Metrics**: Given a proposed config change (new AQM, additional slice, 400 GbE upgrade), the generator produces the expected delay/loss/jitter distributions before any byte is moved, letting operators compare KPI deltas without real-world probing.

Use cases of Agentic AI for Operations & Management (1/3)

- Intent Based Service Delivery

```
+----------------------------------------------------------+
|            L3VPN Service Delivery Application            |
+-------------------------+--------------------------------+
                          |
                   Intent |LPI
                  interface
                          |
+---------+  +------------V--------------------------------+
|         |  |                                             |
|Knowledge|  |       Network Management AI Agent           |
|  Base   <-->                                             |
|         |  |  +----------------------------------------+ |
+---------+  |  |     Intent Decomposing&Analysis        | |
             |  |                                        | |
+---------+  |  | +----------++----------+ +------------+| |
|         |  |  | |  Config  || Config   | |   Config   || |
| Network |  |  | |Generation||Validation| |Distribution|| |
| Digital <-->  | +----------++----------+ +------------+| |
| Twin    |  |  +----------------------------------------+ |
| Tools   |  |                                             |
+---------+  +---------------------------------------------+

+----------------------------------------------------------+
|            Network  Infrastructure                       |
+----------------------------------------------------------+
```

Legend: LPI – Language Programming Interface

Use cases of Agentic AI for Operations & Management (2/3)

- Cross-layer and Cross-domain Multi-Agent communication for Complaint handling

```
                     +------------------------+
                     |Automobile Manufacturer|
                     |    Complaints          |
                     +------------------------+
                                |
                     +-----------------------+
                     | BSS Trouble tickets   |
                     |      System           |
                     +-----------------------+
                                |
              +-------------------------------------------+
              |                  OSS                      |
              |                                           |
              | +---------------+       +---------------+|
              | |   Complaint   |       |   Complaint   ||
              | | Handling Agent|-------| Handling Agent||
              | |  In Domain A  |       |   In Domain B ||
              | +-------+-------+       +---------------+|
              +---------+---------------------------------+
                     +--+---------------------+
                  +----+---------+      +------+-------+
                  |   Mobile     |      |    IP        |
                  |   Network    |      |   Network    |
                  |   Agent      |      |   Agent      |
                  +--------------+      +--------------+
```

Use cases of Agentic AI for Operations & Management (3/3)

- AI Agent Driven Network Management

```
                              +-----------------+
                              |       OSS       |
                              +--------+--------+
                                       |
                                       |LPI
                                       |
                                       |
         Model Invocation    +---------+--------+
  +--------------------------+                  |
  |         Momery Access     Network Management -----------+
  |        +-----------------+     AI Agent     |           |
  |        |            +----|                  |           |
  |        |            |    +-------^---+------+           |
  |        |            |   Response |   |Execution         |
  |   +----V---+   +----+-----+    +-+---V-----+     +------+------+
  |   |        |   |          |    |           |     |             |
  |   | Memory +---|Tools-box <----| Validation+-----> Execution   |
  |   |        |   |          |    |           |     |             |
  |   +--------+   +----^-----+    +-----------+     +-------------+
  |              Model  |     Tool Calling     Action Execution
  |           Invocation|
  |                     |
+-V---------------------V------------------------------------------+
|                     Model Repository                             |
|  (Task Decomposing, Reasoning, Data Analysis, Decision Making..) |
|              +--------------+      +-----------------+           |
|              |  LLM Models  |      |Specialized Small|           |
|              +--------------+      |    AI Models    |           |
|                                    +-----------------+           |
|                                                                  |
+------------------------------------------------------------------+
```

Thanks!!

Questions to NMRG