AI based Network Management Agent (NMA): Concepts & Architecture

NMRG, IETF 121

draft-zhao-nmop-network-management-agent/

zhaoxing@caict.ac.cn

Xing Zhao (CAICT)
Yunbin Xu (CAICT)
Chaode Yu (Huawei)
Haijun Meng (CAICT)
Yipeng Fu (CAICT)

Background

Introducing AI into network management has reached a consensus:

- Enhancing the level of network intelligence and creating Autonomous Networks (AN) [TMF-IG1230] or Autonomic Networking [RFC7575] has become a global consensus among operators, with mainstream operators releasing goals and plans to **achieve Level 4 (L4) autonomous networks by 2025**.
- L4+ AN sets higher requirement in intention, decision-making, analysis, perception, and execution. Artificial Intelligence (AI) technology like ML, LLM has shown significant advantages and great potential in **identification, understanding, decision-making, and generation**, which can well match the new requirements of Level 4 AN and already be one of the core driving technologies to achieve high-level AN.

Key challenges needs to be settled:

- The **application architecture** of AI in network management is not clear.
- The relationship between AI and the existing network management and control systems is not clear.
- **In what form** it can help network management.
- New interface capability requirements after AI is introduced are not clear.

The performance of AI technology itself is what AI technology vendors should consider, and what we need to focus on is how to use it.

Why propose this draft

Related works:

- RFC 7575 Autonomic Networking: Definitions and Design Goals (RFC 7575): *defines autonomic node (device layer)*
- RFC 9315 Intent-Based Networking - Concepts and Definitions: *this draft is the means to implement IB networking*
- Research Challenges in Coupling Artificial Intelligence and Network Management (draft-irtf-nmrg-ai-challenges-03)
- AINEMA (draft-pedro-nmrg-ai-framework-05): *defines AINEMA architecture based on MAPE functions.*

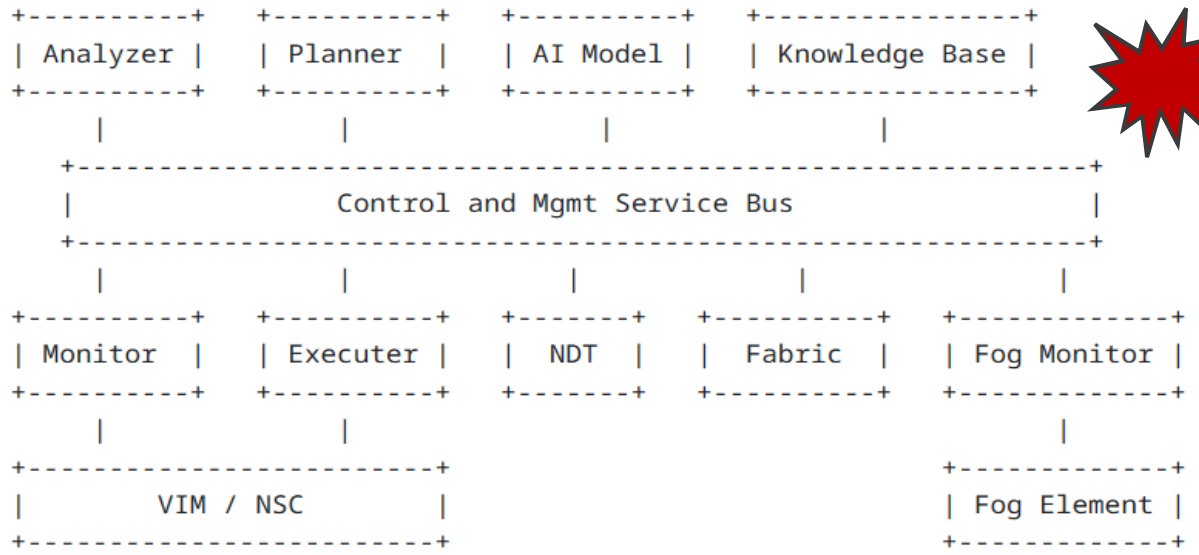


Figure 1: AINEMA Architecture

Some thoughts:

- Analyzing is capability that AI model can provide and probably should not be separated from AI model.
- ML/AI model=**brain**, we need **body (planner, executer, etc)** to **action**.
- Instead of multiple small function modules, we need **an whole entity that provides closed-loop intelligent management**. And it's best to be consistent with the current SDN controller architecture for easy evolution

Agent=AI model+Planning+Action/Execution+Memory

Agent may be one of the most feasible application **form** of AI in network.

Standardize the general framework and capabilities of the Agent, rather than defining how the agent is implemented

The goal and scope of this draft

- This draft is trying to **give a standardized common architecture for the use of AI in network--->network management agent (NMA)**.
  - Provide the framework of **network management agent** (AI agent for network management) Clarify the relationship of NMA with existing MCS or other control systems
  - Define **functional requirements** of agent instance for different scenarios and related interfaces

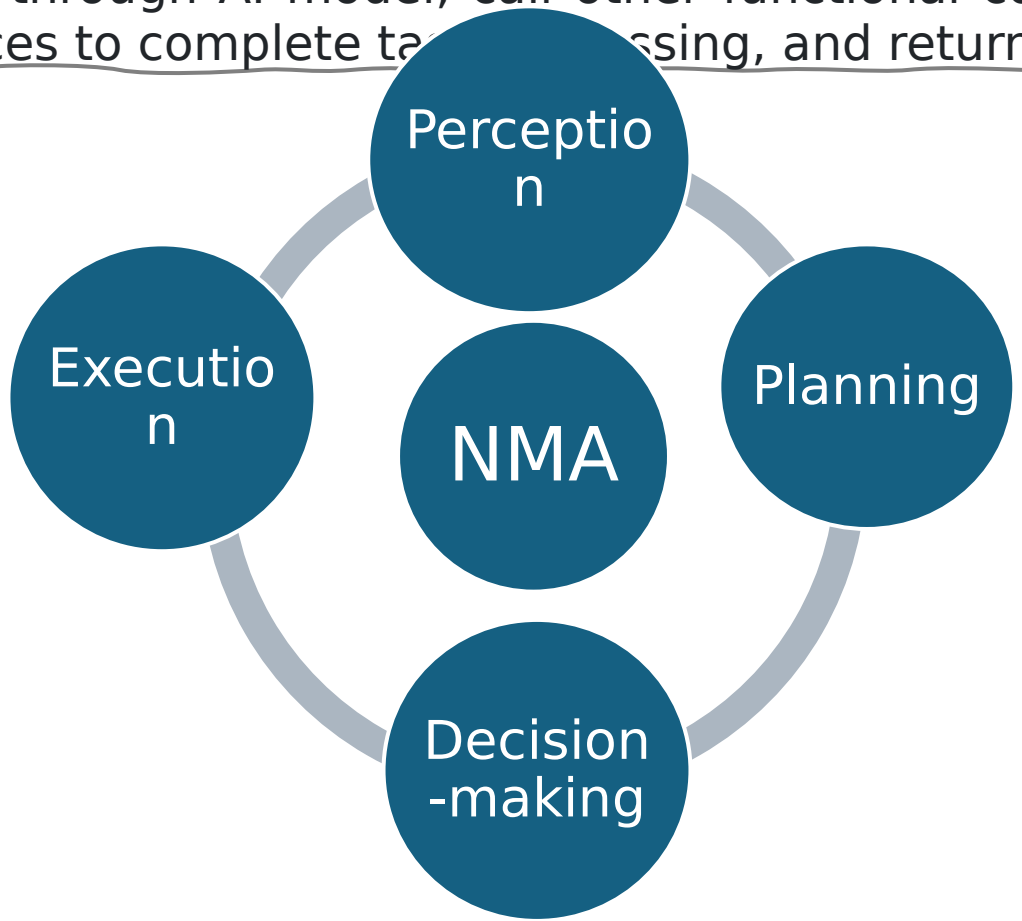
This work is focus on the common framework of control plane.

The specific agent implementation details are not in the scope of this draft.

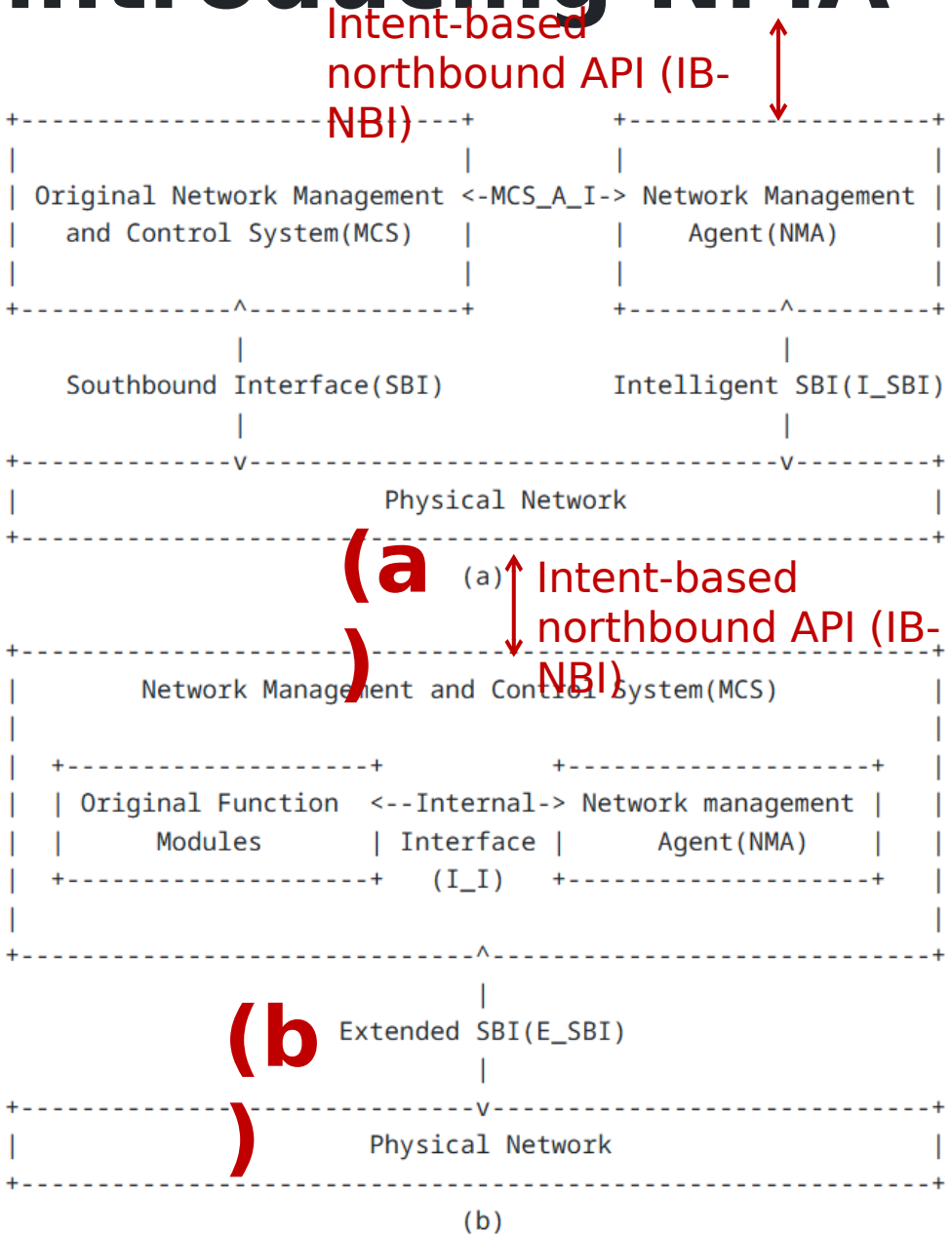
Concepts of NMA

Network management agent (NMA):

- ***A network management entity with autonomous task processing capabilities, which is encapsulated based on AI algorithm or AI models, and has task intent [RFC9315] perception, planning, decision-making, and execution capabilities.*** It can understand the input operation intent through AI model, call other functional components of the control system or external interfaces to complete ta[illegible]ssing, and return processing results.



Open issue (1): Common framework of introducing NMA



A or B?

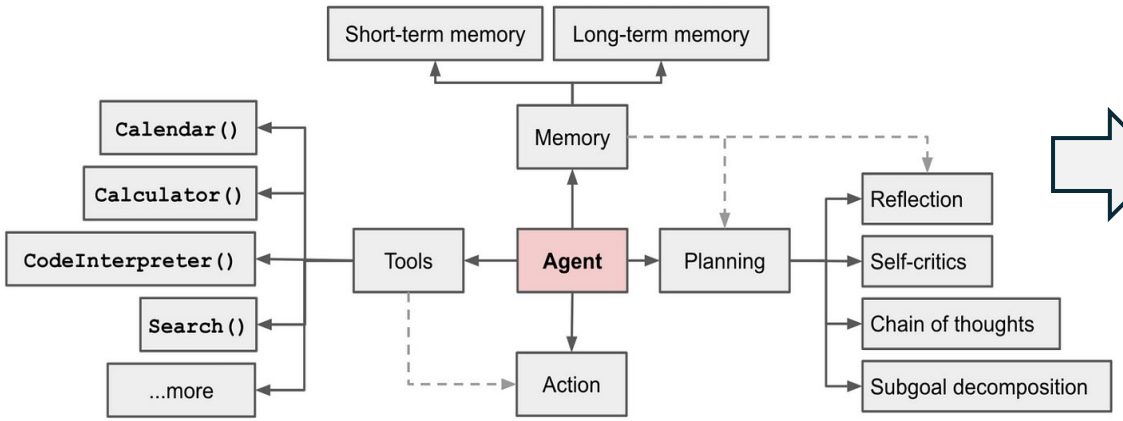
- **A) Independent mode:** NMA is independent from the original network management and control system (MCS).
  - A new east-west interface (MCS_A_I) needs to be added between NMA and MCS to achieve capability calling and result feedback.
  - An intelligent southbound interface (“I_SBI”) needs to be added between NMA and the physical network.

- **B) Integrated mode:** NMA is integrated with MCS, and the NMA serves as a function of MCS.
  - NMA interacts with original function modules through internal interface (“I_I”).
  - The enhanced MCS interacts with underlay physical network through extended SBI (“E_SBI”)

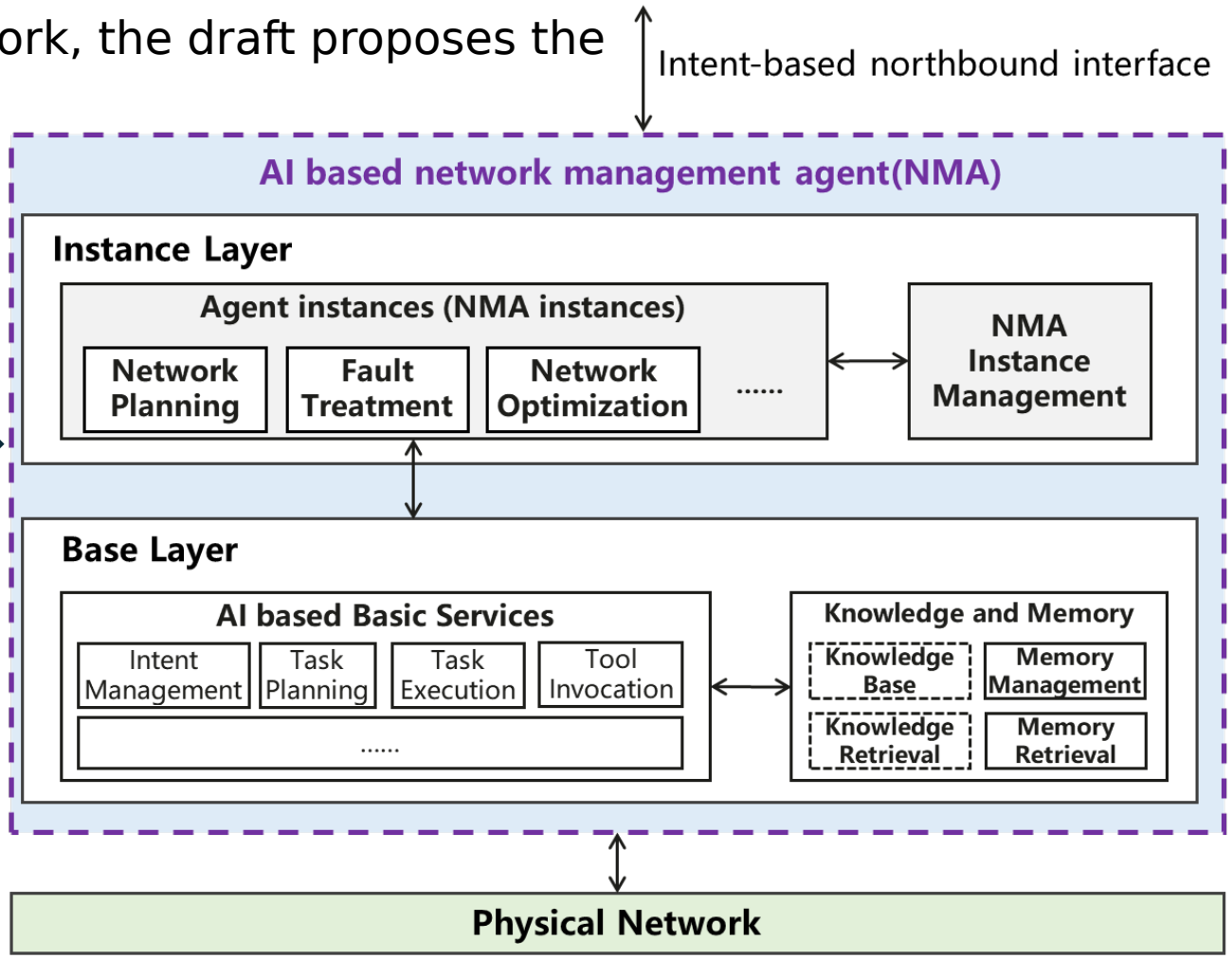
In both modes, it is necessary to add a intent-based northbound API (IB-NBI), only the capabilities of the API and the information that needs to be exchanged differ between the two modes.

Reference function architecture of NMA

By referring to the common AI agent framework, the draft proposes the reference function architecture of NMA.



Common AI Agent framework



Reference network management Agent architecture

Call for feedbacks!

Some concepts in the reference architecture of NMA

NMA instance:

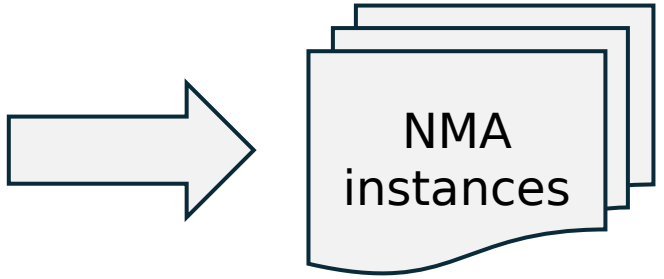
- The instantiated agent applications which can automatically perform certain network management tasks for specific network management scenarios. For different scenarios, there can be multiple scenario-oriented agent instances (*like apps in the phone*), which can be called "*NMA instance*" for short in this document. For different application scenarios, there can be multiple scenario-oriented NMA instances.

Network faults handling

Network optimization

Service configuration

.....



NMA instances in NMA are like apps in the phone

Some concepts in the reference architecture of NMA

1) Base layer:

- **AI based basic services**: Provide *a unified intelligent agent engine framework*, build interactive intelligence public capabilities, simplify instance development, integrate Large Language Models (LLM), knowledge retrieval, API invocations, etc., to *achieve the full process orchestration from intent understanding, task planning, tool invocation to task execution*.
- **knowledge and memory subsystem:**
  - Provides unified search for local multi-type knowledge bases
  - combines LLM to complete knowledge fusion and extraction
  - improves the accuracy of downstream tasks
  - realizes knowledge injection and integrated retrieval.

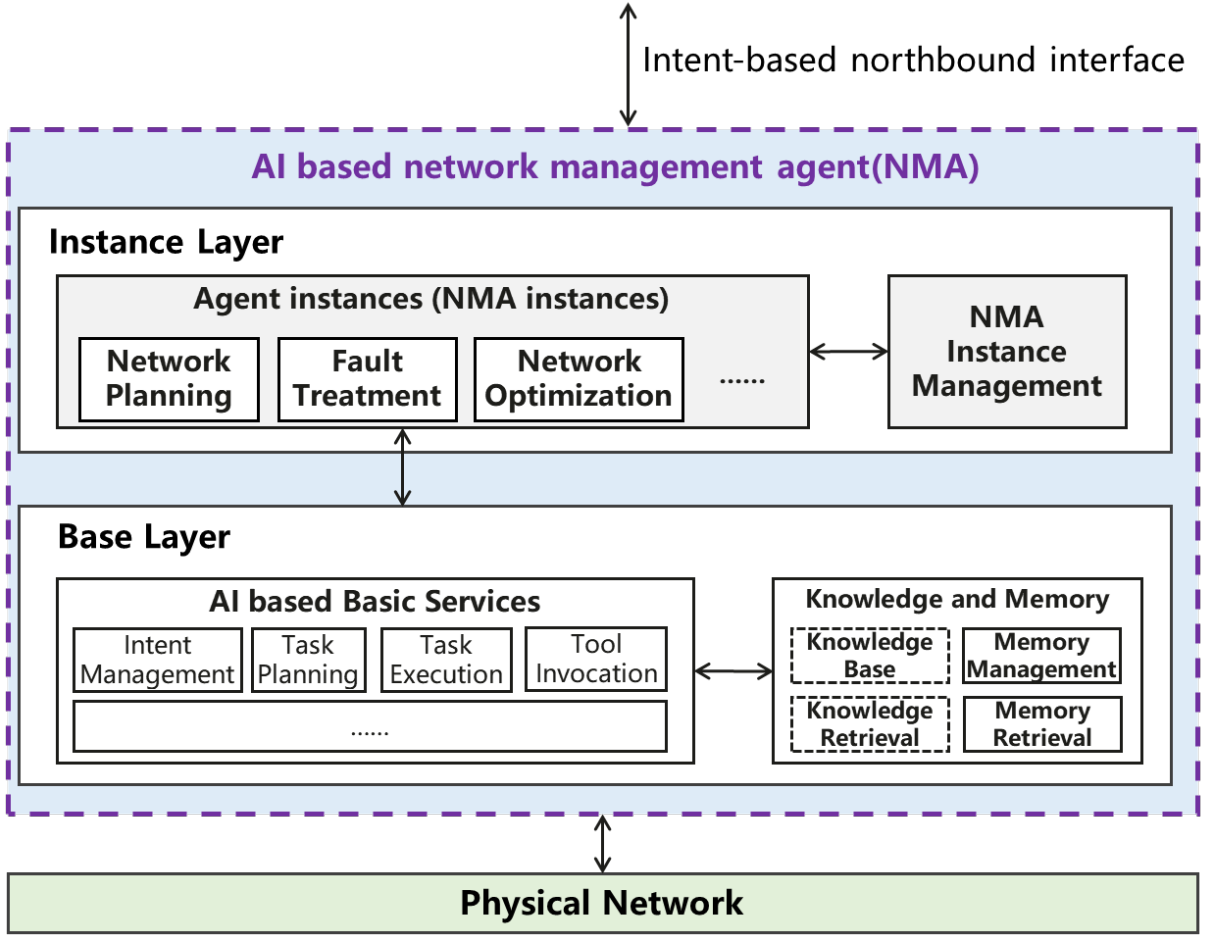
2) Instance layer:

- **Agent instances (NMA instances)**: Aimed at the network planning, construction, maintenance, optimization, and operation scenarios, the main NMA instances could include:
  - Network Fault Handling Instance
  - Network Planning Instance
  - Network Optimization Instance
  - Intelligent Assistant Instance, etc.
- **Instance management:** Implements basic management capabilities including:
  - registration, lifecycle management, operation monitoring, and log auditing, etc.

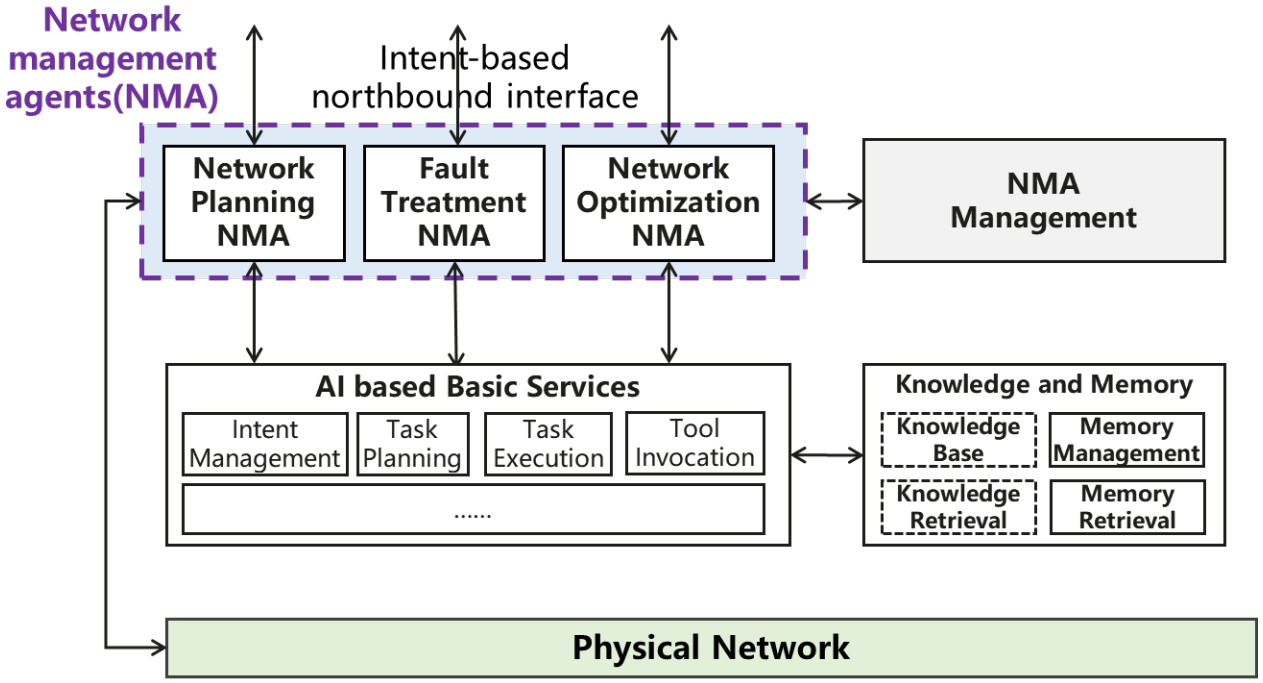
Open issue (2): Architecture of NMA

A or B?

Option a):
AI capability as the *internal* base layer of NMA, with multiple *NMA instances* for various scenarios

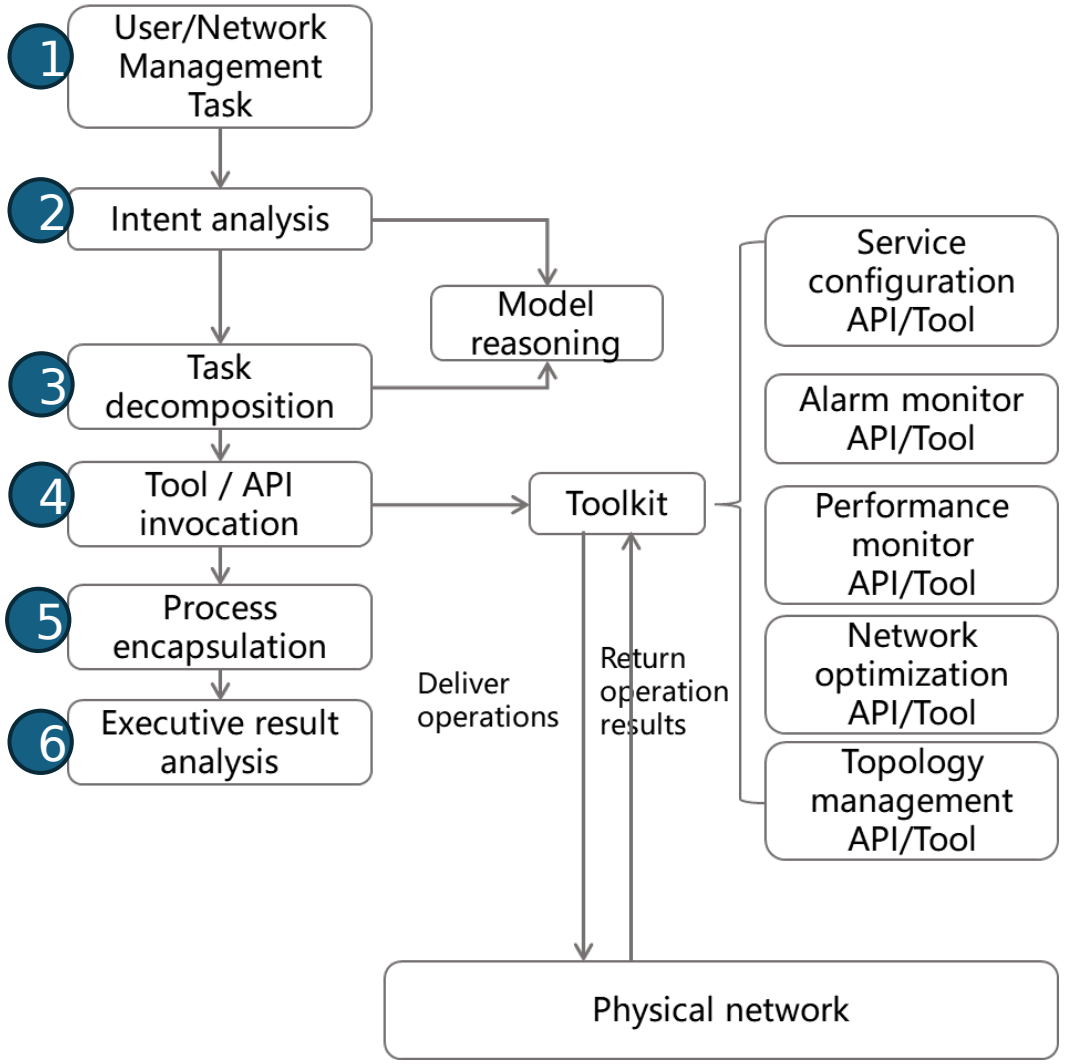


Option b):
AI capability as the *external* service of NMA, with individual *NMA* for each scenario



The two options may affect the implement of NMA.
We currently prefer the first approach, in which we can extend *NMA instances* with plug-ins instead of deploying new NMAs one by one.

Common workflow of NMA



① **User/Network Management Task Input:** Input the user's task information through multiple rounds of natural language interaction.

② **Intent Analysis**: Analysis user task intent through AI model reasoning provided by the AI based basic services within NMA.

③ **Task Decomposition**: **Split the task into detailed operations** to be performed based on the analyzed intent of the task.

④ **Tool/API Invocation**: Call the corresponding tool or function API to complete the execution of each operation listed in step 3). The ***toolkit*** refers to the collection of all tools that can be used directly to manage and operate physical networks.

⑤ **Process Encapsulation**: Package the operation results of all steps into the execution result of the entire task.

⑥ **Executive result analysis**: Analyze the task processing results and return to the user.

Goal: Achieve closed-loop automated processing of tasks and provide end-to-end intelligent network maintenance assistance.

Typical scenarios

① **Network management and maintenance scenarios, including:**

- **Intelligent planning and construction**: such as broadband installation, resource/capacity planning, intelligent acceptance, site selection, etc.
- **Intelligent maintenance**: such as intelligent fault diagnosis, quality analysis, operation and maintenance/cutting assistant, broadband maintenance assistant, etc.
- **Intelligent optimization**: such as route optimization, coverage optimization, topology optimization, and intelligent energy saving, etc.

② **Network operation scenarios, including:**

- intelligent question and answer, customer service assistant, product recommend, intelligent marketing and other value-added services. *This part is outside the scope of this document.*

- ***Application considerations**: strong demand, feasible technology, and good input-output ratio, sufficient data for AI pre-training, perfect data annotations, high fault tolerance rate.*
- ***First possible mature application scenarios**: broadband installation and maintenance assistant, fault diagnosis, operation and maintenance assistant, etc.*

Next steps

1. Call for feedbacks on the **open issues**, zhaoxing@caict.ac.cn
2. Clarify the **function architecture** of NMA, decide whether base layer belongs to NMA.
3. Specify the **NMA/ NMA instance functional requirements** for different scenarios.
4. Discuss the **functional requirements and definitions for each interface**.
    - Discuss whether XML/YANG adapts to the interface interaction requirements of AI intelligent operation, If yes, ***IB-NBI YANG model can be defined further*** *in NMOP, CCAMP or other groups.*

Thanks for listening !

zhaoxing@caict.ac.cn