



Introducing Generative AI Solution Based on Large Models

Lingli Deng (CMCC), Keguang He (CMCC)

Contents

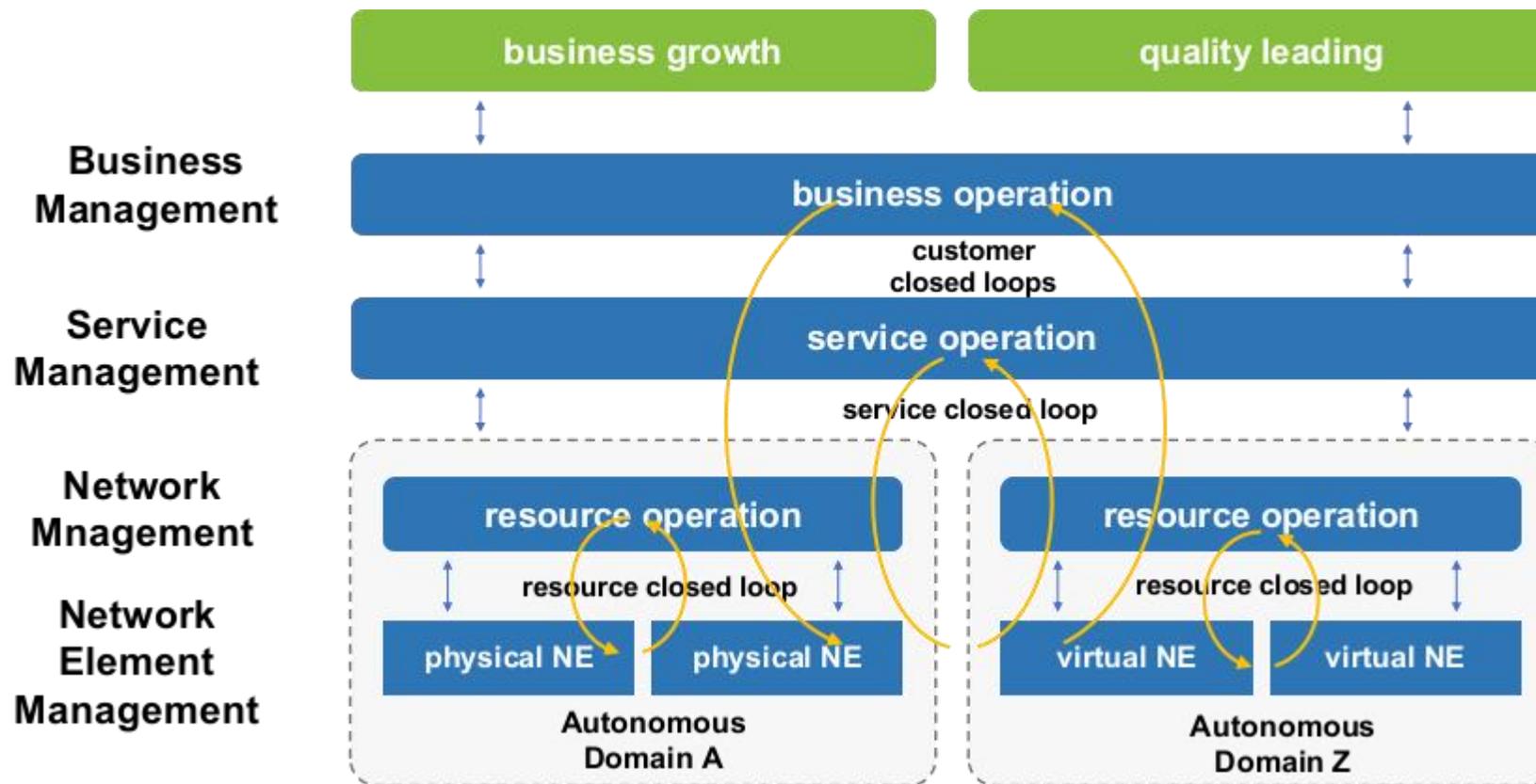
01 **Motivation**

02 Requirements Introduction

03 Project Impact

Ref: ITU-T "Scenarios and Requirements of Intent-Based Network for Network Evolution"

Autonomous Network System Architecture



Autonomous network: A network composed of intelligent and automated infrastructure, operational management systems, and business systems.

Challenges in the Evolution of Autonomous Network



- There is currently no unified consensus on the technical paths and evolution strategies for advancing towards high-level autonomous networks.
- Key technologies such as network AI large models and generative AI urgently need to accelerate breakthroughs
- How to establish an open and cooperative ecosystem, fostering industry chain synergy to achieve scale advantages, is a critical challenge that telecommunication operators or industry organizations need to focus on during top-level design.

Key Enabling Technology

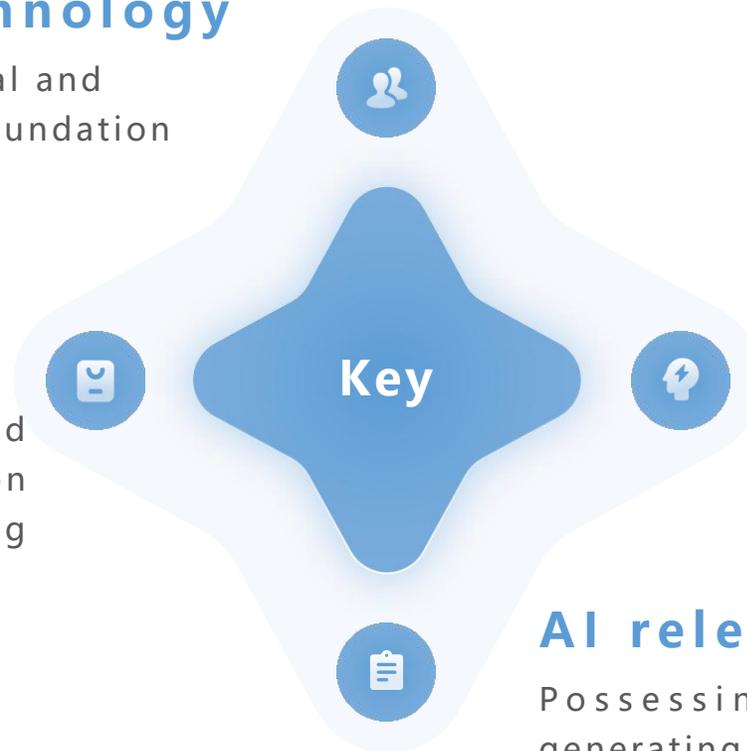
To achieve the goal of reaching Level 4 advanced autonomous networks by 2025, there is a need to focus on exploring the applications of key technologies such as network convergence awareness, network digital twins, intent-driven, generative AI, and others. Through technological innovation, to realize a significant leap in application value.

Network Convergence Awareness Technology

The integration and perception of internal and external network environments are the foundation for achieving autonomous network.

Network Digital Twin Technology

Deliver comprehensive, customizable, and iterative business pre-validation capabilities with network-wide monitoring



Intent-Driven

Key to Enhancing Network Intelligence

AI related technology

Possessing general understanding capabilities, generating diversity, and featuring boundless knowledge, it serves industry-specific applications.

Elevating network intelligence through AI-based intent automation

Intent and AI complement each other in autonomous network solutions. AI will replace/supplement traditional algorithms, addressing the complex optimization tasks faced by intent. And intent will enhance the use of AI and sometimes become the foundation for training AI models.

Intent Recognition and Parsing

Generative AI can assist the system in more accurately identifying and parsing user intent and requirements.

Personalized Service Customization

By learning and analyzing users' historical data and behavior patterns, Generative AI can provide personalized network feature configuration solutions for different users



Intelligently Matching and Configuring Features

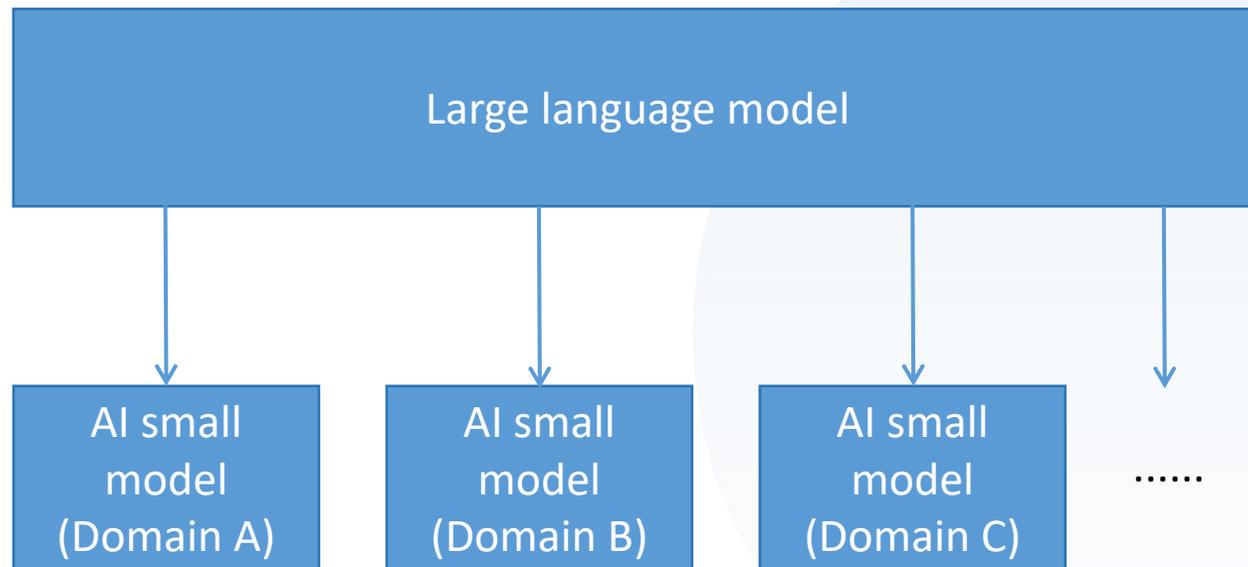
Generative AI can intelligently match appropriate network features and perform corresponding configurations based on user intent.

Context Awareness and Selfadaptation

Recognizing and understanding changes in user intent across different contexts, and intelligently adjusting the deployment and configuration

AI introduces entirely new capabilities

By harnessing the unique advantages of new technologies such as generative AI, including 'network panoramic knowledge cognition,' 'emergence of new scenario capabilities,' and 'intent-driven interactions,' it will drive the evolution of traditional networks towards 'intent-driven' networks, bringing a completely new impetus to the digital transformation and upgrading of networks.



- The AI small models and API capabilities, functioning as a collective, are encapsulated with a large language model serving as an orchestration scheduler at the outer layer.
- These AI capabilities span various areas, including core networks, wireless, government-enterprise solutions, and home broadband, covering a diverse range of scenarios such as perception, diagnosis, control, and prediction.
- Based on user input, the large language model, through understanding user intent, invokes corresponding AI small models and APIs or their combinations.
- Leveraging the capabilities of a multitude of AI small models and APIs, it enhances operational efficiency, with the large language model serving as an interface to provide unified services internally and externally.

Contents

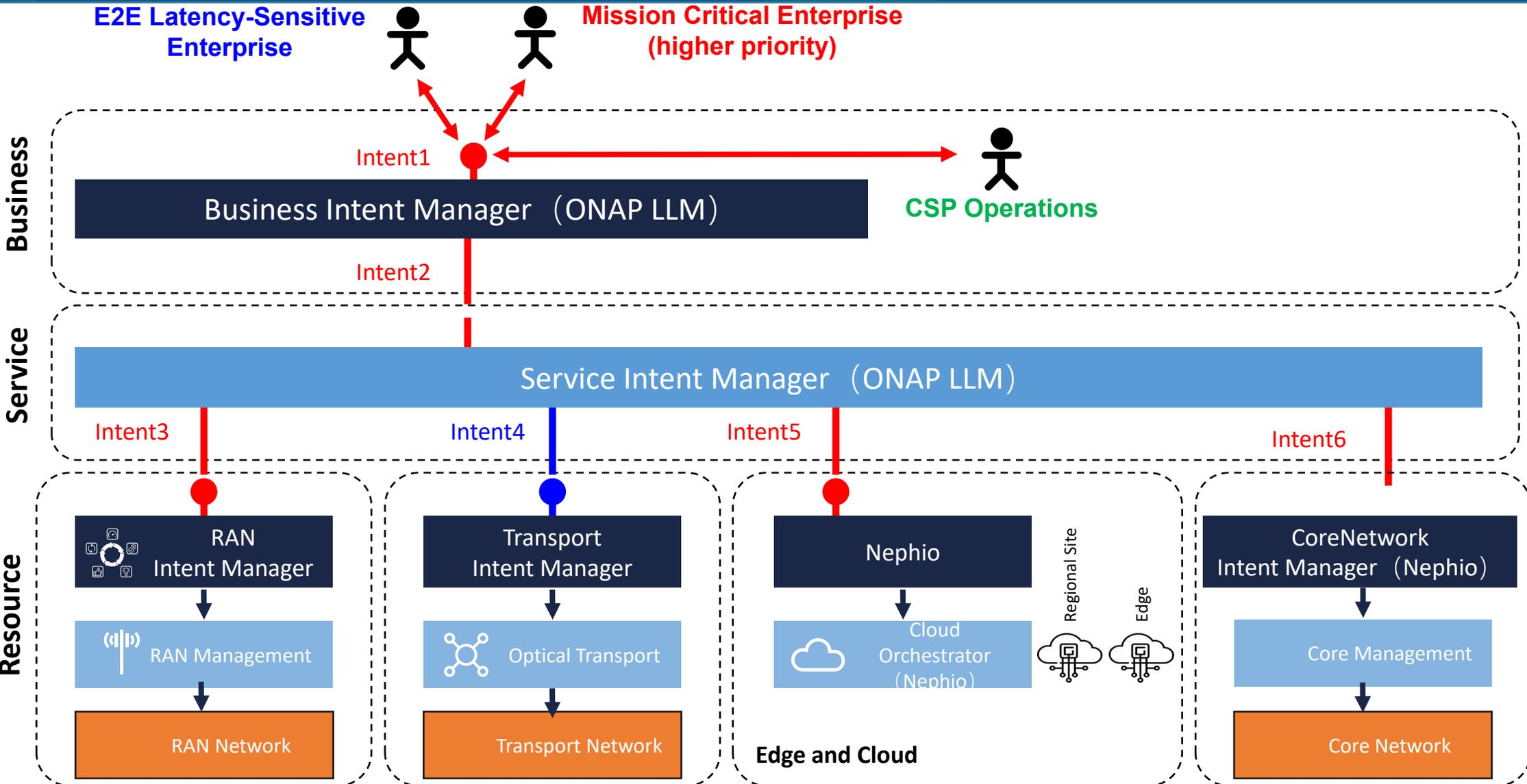
01 Motivation

02 **Requirements Introduction**

03 Project Impact

Ref: ITU-T “Scenarios and Requirements of Intent-Based Network for Network Evolution”

Enhancing Autonomous Network through Generative AI and Intent-Driven Technologies in ONAP



Proposed Projects

- **Business Intent Manager & Service Intent Manager: ONAP**

- 1、 REQ-1267: Provide a general intent model and a general intent interface to ensure that all intents (especially machine - machine intents) operate according to the same expression and process.

- 2、 Proposing project initiative for the ONAP New Delhi Release: Introduction of Generative AI Capabilities Based on Large Models in ONAP

- **Core Network Intent Manager: Nephio**

Nephio has officially launched its first release, making it a highly acclaimed project within the LFN community. In this first release, use cases pertaining to UPF (User Plane Function) and SMF (Session Management Function) have been included. These cases are instrumental for showcasing the capabilities outlined in our 5G Service Blueprint (SBP) project.

- **RAN Intent Manager, Transport Intent Manager, Edge Manager (Infosys EMCO) :**

Waiting for further discussion. Any suggestions?

Proposed Assignment (TBD)

(1) Business Layer:

- Intent Management: **ONAP UI module -- China Mobile**

(2) Service Layer:

- Intent Management: **ONAP UI module -- China Mobile**
- Service Orchestration: **To Be Added**

(3) Resource Layer:

- RAN Intent Management: **To Be Added**
- RAN Management: **To Be Added**
- RAN Network: **To Be Added**

- Transport Intent Management: **To Be Added**
- Transport Management: **To Be Added**
- Transport Network: **To Be Added**

- Core Intent Management: **Nephio -- China Mobile, Infosys**
- Core Management: **To Be Added**
- Core Network: **To Be Added**

- Edge Application Intent Management: **To Be Added**
- Cloud Orchestrator: **To Be Added**
- Edge Resources: **To Be Added**

REQ-XXX Introducing Generative AI Solution Based on Large Models

In autonomous network, ONAP can provide capabilities for network orchestration and management. Introducing AI capabilities based on large models can empower ONAP with more intelligent decision-making, optimization abilities, and an enhanced customer experience. Therefore, we can incorporate various large models into ONAP and provide further integration capabilities with its existing functionalities.

Key Contacts - Lingli Deng(CMCC), Keguang He (CMCC)

Executive Summary - Integrating AI capabilities into ONAP enhances network orchestration with intelligent decision-making and optimization, elevating the customer experience and reinforcing ONAP's position in autonomous network.

Business Impact - The infusion of large models into ONAP translates into tangible benefits for businesses. Enhanced network orchestration through intelligent decision-making and optimization not only boosts operational efficiency but also delivers a superior customer experience.

Business Markets - This REQ provides a novel solution for advancing network orchestration capabilities by introducing generative AI solution based on Large Models into ONAP. Industries such as telecommunications, IT, and network services are poised to benefit from enhanced operational efficiency and improved service delivery.

Funding/Financial Impacts - The integration of introducing generative AI solution based on Large Models into ONAP represents a noteworthy investment with promising financial returns, the long-term benefits include increased operational efficiency, reduced maintenance costs, and improved service quality.

Organization Mgmt, Sales Strategies - *There is no additional organizational management or sales strategies for this requirement outside of a service providers "normal" ONAP deployment and its attendant organizational resources from a service provider.*

Why ONAP need LLM

More comprehensive user interaction: Users can interact with the system in a more natural way, providing high-level instructions and requests rather than just predefined commands.

Intelligent Decision Support: Large models can analyze complex network data and events to support intelligent decision-making. For example, they can be used to identify abnormal behavior in the network, predict potential issues, and offer optimization suggestions.

Task Automation Description and Generation: Large models can be employed to generate descriptions for automated tasks, aiding users in creating and understanding automation scripts more easily.

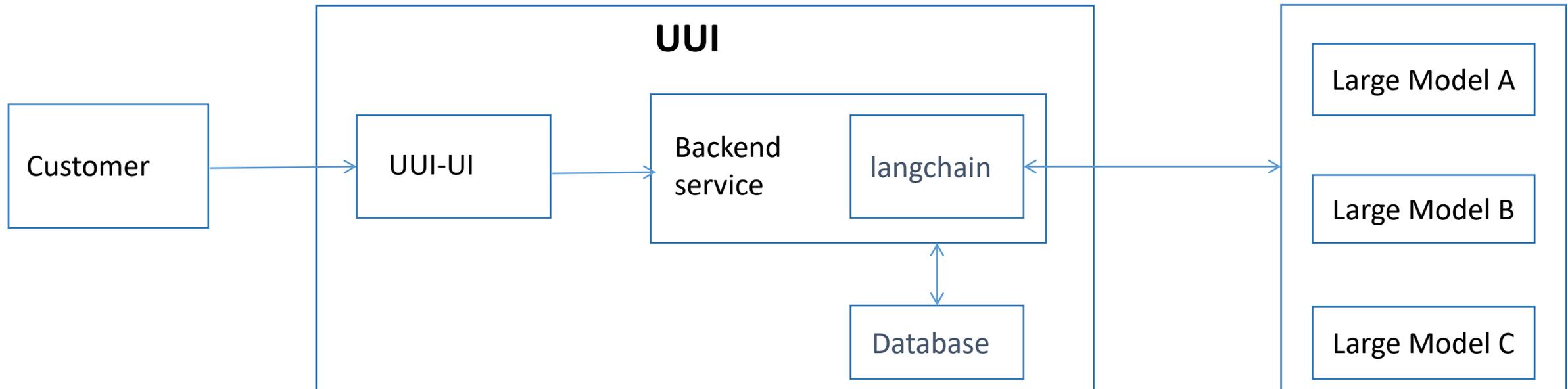
Log Analysis and Fault Diagnosis: Large models can analyze extensive network logs, identifying potential issues and faults.

User Support and Training: Introducing large models in ONAP can enhance user support and training processes. Models can automatically respond to user queries, provide training materials, and interact with users in a more intelligent manner.

UUI incorporates LangChain to introduce the capability of large models

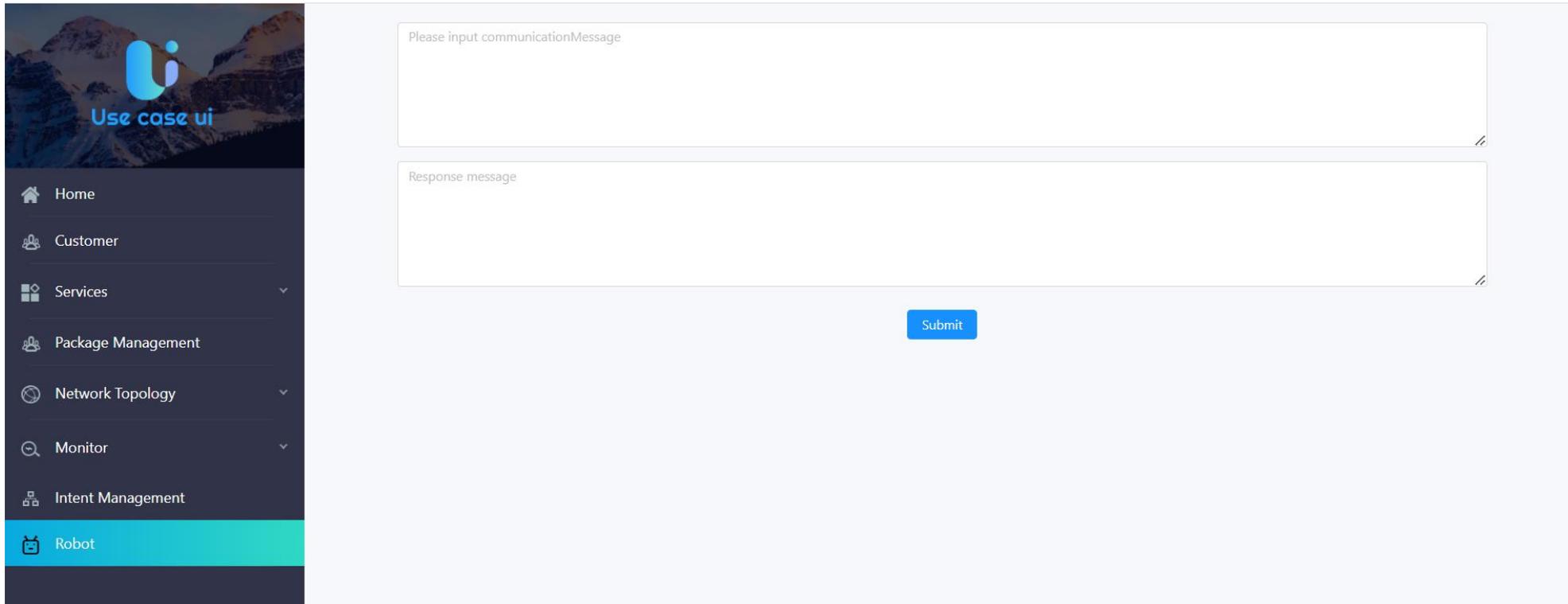
We are not developing the capability for large models, but rather integrating existing industry-leading large models through UUI.

- The backend services of UUI can integrate with LangChain, offering the capability to handle large models.
- LangChain allows integration with multiple large models, enabling customer to choose and integrate different large models through UUI based on their requirements.
- UUI-UI provides customers with a unified input interface. The database is used to store usage records of large models for optimizing user experience



Provide a Unified Frontend Interface

- Providers a Q&A Page
- Record historical execution information on the frontend page
- Provide the capability to select large models



The screenshot displays a web application interface. On the left is a dark sidebar menu with a teal header containing a logo and the text "Use case ui". The menu items are: Home, Customer, Services, Package Management, Network Topology, Monitor, Intent Management, and Robot (highlighted in teal). The main content area is light blue and contains a Q&A form. The form has two text input fields: the top one is labeled "Please input communicationMessage" and the bottom one is labeled "Response message". Below the input fields is a blue "Submit" button.

Provide functionality for evaluating large models

- UUI can integrate open-source tools for evaluating large models, such as AI-Benchmark:
- Install the large model evaluation tool, for example, AI-Benchmark.
- Prepare a testing dataset; AI-Benchmark supports loading datasets from various sources.
- Create an evaluation script, configuring the path and parameters of the large model to be evaluated.
- Run the evaluation task: Within the evaluation script, invoke the AI-Benchmark evaluation API to execute the evaluation task and obtain corresponding results and metrics.
- Analyze the evaluation results: Based on the results and metrics, performance and capabilities of the large model can be analyzed and assessed. The evaluation results can be presented to users in the form of a report.

Possible supported functions

In the process of intent processing, large models can be used to accomplish many functions.

- Func1: Introduce large models to achieve intelligent question answering functionality.
- Func2: Introduce large models to achieve natural language processing and convert it into formatted intents.
- Func3: Introduce large models to convert formatted intents into specific operational configurations or strategies.
- Func4: Utilize the simulation platform to integrate intent verification functionality.

Scope

- **MVP tasks include:**

1. Provide generative AI interactive functionality based on large models.
2. Supply a large model adaptation layer that can integrate with different large model services through configuration.
3. Deliver large model evaluation functionality to assess different large models.
4. Provide end-to-end use cases for the utilization and testing of large models.

- **Potential extensions include:**

1. Integrate generative AI and large models into the intent processing use cases in the existing ONAP requirement REQ-1267 "General intent model and general intent interface requirements"

Welcome everyone to join us.



ONAP

OPEN NETWORK AUTOMATION PLATFORM

Thanks!

hekeguang@chinamobile.com