DEMO: Next-Generation 6G Network Management with OSS-GPT

Abdelkader Mekrache
EURECOM
France
abdelkader.mekrache@eurecom.fr

Adlen Ksentini
EURECOM
France
adlen.ksentini@eurecom.fr

Christos Verikoukis
ISI/ATH and University of Patras
Greece
cveri@isi.gr

ABSTRACT

In the 6G era, the Management and Orchestration (MANO) layer must become more intelligent to support advanced use cases with stringent Quality of Service (QoS) requirements. These often span multiple domains, terrestrial, non-terrestrial, and edge, requiring efficient coordination across heterogeneous networks. To meet these demands, Mobile Network Operators (MNOs) are evolving their Operations Support Systems (OSS) toward a "network of networks" vision, built on standardized and interoperable APIs. Standards bodies such as 3GPP, ETSI, and TM Forum are addressing this by developing numerous OSS API specifications and intent-based request models in JSON or YAML. However, creating these intents remains challenging for users without deep domain knowledge. We present OSS-GPT, an agentic AI framework that enables users to interact with OSSs via natural language. By leveraging Large Language Models (LLMs), OSS-GPT translates high-level intents into executable sequences of OSS API calls and automates their execution.

CCS CONCEPTS

• **Networks** → **Network management**; • **Computing methodologies** → **Natural language processing**.

KEYWORDS

MANO, MNO, OSS, IBN, API, LLMs.

ACM Reference Format:
Abdelkader Mekrache, Adlen Ksentini, and Christos Verikoukis. 2025. DEMO: Next-Generation 6G Network Management with OSS-GPT. In *ACM SIGCOMM 2025 Conference (SIGCOMM Posters and Demos '25), September 8–11, 2025, Coimbra, Portugal.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3744969.3748429

1 INTRODUCTION

6G networks are expected to deliver advanced services with stringent Quality of Service (QoS) requirements. To meet these demands, Mobile Network Operators (MNOs) are deploying advanced AI-powered Operations Support Systems (OSS) to efficiently manage their infrastructures and 6G services. This approach is critical for realizing the "network of networks" concept, which relies on intelligent interoperability among the OSS platforms of different MNOs. To facilitate this, standardization bodies such as 3GPP, TMForum, and ETSI are actively defining APIs for OSSs to facilitate interactions, which include declarative structures, allowing OSS users (i.e., MNOs administrator, Network Operations Center (NOC) engineers) to declare their intentions, thus enabling the concept of Intent-Based Networking (IBN).

However, learning these declarative structures of IBN, which rely on YAML and JSON formats, is not straightforward, especially for users with limited domain knowledge. Consequently, researchers are focusing on enabling natural language interaction with OSS. In this approach, Large Language Models (LLMs) are used to translate it into standardized structures. However, existing research often focuses on simple intents that tackle only a single objective, such as defining the parameters of one API call in natural language [4]. In real-world scenarios, a high-level intent may require multiple OSS API calls, which must be executed sequentially.

In this demo, we present OSS-GPT, a novel 6G IBN framework that enables zero-touch OSS management via natural language. Unlike static intent-mapping systems, OSS-GPT uses an agentic LLM architecture to dynamically interpret user intents, decompose them into OSS API calls, and execute them. This agentic design marks a step toward fully autonomous network operations.

2 OSS-GPT

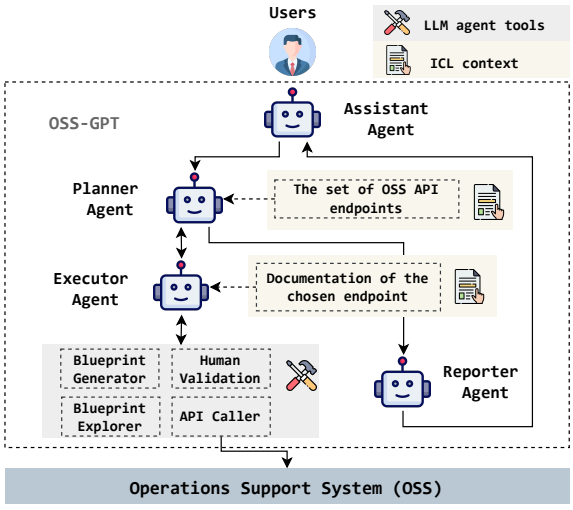

Figure 1: OSS-GPT agentic design.

OSS-GPT follows a multi-agent LLM architecture, inspired by research showing that decomposing complex tasks into role-specific agents improves reliability [3]. Fig. 1 illustrates the agentic workflow, composed of four main agents:

- Assistant: Interface agent that handles user interaction. It forwards actionable OSS intents to the Planner and relays final results returned by the Reporter.

SIGCOMM Posters and Demos '25, September 8–11, 2025, Coimbra, Portugal

ACM ISBN 979-8-4007-2026-0/2025/09.
https://doi.org/10.1145/3744969.3748429

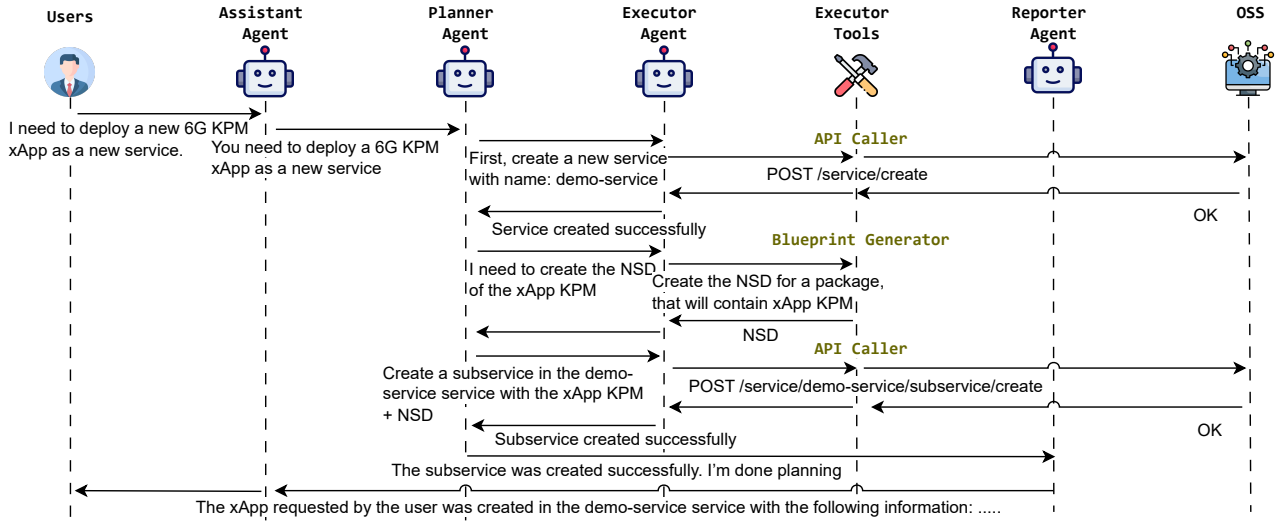

Figure 2: An OSS-GPT use case workflow in which a user deploys a RAN xApp using natural language.

- Planner: Decomposes high-level natural language intents into sequences of API calls. It sends one call at a time to the Executor and replans if errors occur. After execution, it requests a natural language summary from the Reporter.
- Executor: Executes API calls by parsing parameters from the specification and using internal tools: a Blueprint Generator (e.g., to generate manifests), Blueprint Explorer, and API Caller. It also supports Human Validation before critical operations (e.g., POST/DELETE).
- Reporter: Summarizes the execution sequence in natural language based on the Planner–Executor exchange and returns the summary to the Assistant for user delivery. This enables natural language-based intent reporting in IBN systems.

For more details, readers can refer to [5].

3 DEMONSTRATION

3.1 Demo Setup

We demonstrate OSS-GPT on the EURECOM OSS [1], a standards-aligned system that manages end-to-end 6G services. The service instantiation interface complies with ETSI standards and uses Network Service Descriptors (NSDs) to define applications and their resource requirements. Most OSS-GPT agents (i.e., *Assistant*, *Planner*, *Executor*, and *Reporter*) rely on OpenAI's GPT-4. However, for the blueprint generator, which produces NSDs from natural language, we developed a custom LLM called *NSD-expert*, based on Llama 3.2[1]. To train it, we created a dataset from prior experiments [4] and augmented it using a general-purpose LLM, i.e., GPT-4, via few-shot learning. We then fine-tuned Llama 3.2 with LoRA [2] on the resulting dataset. This new LLM, i.e., NSD-expert, is essential for OSS-GPT, as it is used by the Executor whenever NSDs (manifests) need to be generated prior to invoking the API endpoints.

[1]https://huggingface.co/meta-llama/Llama-3.2-3B

The demo runs across two machines. The first hosts a Kubernetes cluster, managing virtualized 6G services. These include OpenAir-Interface (OAI)[2] RAN, OAI Core Network (CN), and edge applications. The second runs the EURECOM's OSS [1] and OSS-GPT [5]. OSS-GPT is built with LangGraph[3] and interacts with both GPT-4[4] (remote) and *NSD-expert* (local, via Ollama[5]).

3.2 Demo Scenario and Results

We simulate a NOC engineer interacting with OSS-GPT using natural language. The user can ask questions (i.e., intents) about network status or deploy services such as a RAN xApp. Fig. 2 shows how an intent is processed: the *Planner* generates API calls; the *Executor* executes them; and the *Reporter* returns a summary to the *Assistant*. Example intents include: What infrastructures are available?, Can you create a gNB? How many core networks are deployed? ...etc. The video is available at: https://youtu.be/A1tTyHhyT80

4 CONCLUSION

This demo presented OSS-GPT, a framework that enables users to interact with OSSs using natural language. By leveraging multi-agent LLMs, OSS-GPT translates high-level intents into actionable API calls, simplifying network management for users without domain expertise. Our demo shows the potential of natural language-driven OSS management in 6G networks, providing a foundation for future research into more adaptive and autonomous network operations.

ACKNOWLEDGMENT

This work is supported by the European Union's Horizon Program under the SUNRISE-6G project (Grant No. 101139257).

[2]https://openairinterface.org
[3]https://www.langchain.com/langgraph
[4]gpt-4o-2024-11-20
[5]https://ollama.com

REFERENCES

[1] Sagar Arora, Karim Boutiba, Mohamed Mekki, and Adlen Ksentini. 2022. A 5G Facility for Trialing and Testing Vertical Services and Applications. *IEEE Internet of Things Magazine* 5, 4 (2022), 150–155.

[2] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021).

[3] Junyou Li, Qin Zhang, Yangbin Yu, Qiang Fu, and Deheng Ye. 2024. More agents is all you need. *arXiv preprint arXiv:2402.05120* (2024).

[4] Abdelkader Mekrache, Adlen Ksentini, and Christos Verikoukis. 2024. Intent-based management of next-generation networks: An LLM-centric approach. *Ieee Network* (2024).

[5] Abdelkader Mekrache, Adlen Ksentini, and Christos Verikoukis. 2025. OSS-GPT: An LLM-Powered Intent-Driven Operations Support System for 6G Networks. In *2025 IEEE 11th International Conference on Network Softwarization (NetSoft)*. IEEE.