



NATIVE OBSERVABILITY FOR PRIVATE 5G: A REFERENCE ARCHITECTURE ON COMBINATION OF NWDAF, SMO, AND EDGE ANALYTICS

Bhaskara Raju Rallabandi

Principal & Chief Technology Advisor, Invences Inc.,

Frisco, Texas, USA

e-mail: techie.bhaskar@gmail.com

ABSTRACT:

The transition to private 5G networks introduces unprecedented complexity in performance monitoring and assurance, demanding an integrated, data-driven observability framework. This paper proposes a novel reference architecture for native observability in private 5G, moving beyond fragmented tools to a cohesive, closed-loop system. The architecture strategically combines three core components: the 3GPP-standardized Network Data Analytics Function (NWDAF) for core-network intelligence and subscriber-aware analytics; the Service Management and Orchestration (SMO) framework for RAN-centric control and automation; and decentralized Edge Analytics platforms for ultra-low latency, application-specific telemetry processing at the network periphery. We detail how these components interoperate through open interfaces to create a unified data fabric, collecting, correlating, and analyzing metrics from the core, RAN, and edge applications. The proposed architecture enables predictive assurance, automated troubleshooting, and intent-based optimization, providing operators with a holistic, real-time view essential for maintaining stringent service-level agreements and operational efficiency in dedicated enterprise environments.

Keywords: *Native Observability, 5G Private Networks, NWDAF, Service Management and Orchestration, Edge Analytics, Reference Architecture.*

I. INTRODUCTION

Private 5G networks are emerging as a critical infrastructure for Industry 4.0, enabling transformative applications like autonomous mobile robots, augmented reality, and real-time process control. These use cases demand stringent performance guarantees, ultra-low latency, and absolute reliability. Traditional network monitoring, reliant on external probes and siloed management tools, falls short in this new paradigm. It provides a fragmented, reactive view, unable to correlate issues across the complex, end-to-end service chain spanning the user equipment, Radio Access Network (RAN), mobile core, and edge

applications. This visibility gap creates significant operational risk, jeopardizing the service-level agreements (SLAs) that underpin the business case for private 5G. To overcome this, a paradigm shift towards native observability is essential. This concept embeds intelligence and data collection directly into the network fabric, providing a holistic, real-time, and predictive understanding of network and service health. Achieving this requires a unified architectural approach that strategically integrates standardized, cloud-native components. This paper proposes a reference architecture that orchestrates the 3GPP Network Data Analytics Function (NWDAF) for subscriber and core-network analytics, the O-RAN Service Management and Orchestration (SMO) framework for RAN intelligence and automation, and distributed Edge Analytics for application-aware, low-latency telemetry processing at the network periphery. By defining the interfaces and data flows between these elements, we present a blueprint for a cohesive observability plane. This integrated system enables closed-loop automation, from predictive anomaly detection and root-cause analysis to intent-based performance optimization, ensuring private 5G networks are not just connected, but truly observable, manageable, and autonomously resilient.

II. LITERATURE SURVEY

The critical importance of observability in modern, software-defined networks is a well-established research domain, with foundational principles emphasizing the need for comprehensive telemetry, correlation, and analytics to manage complexity. Within the specific context of 5G, the 3GPP's introduction of the Network Data Analytics Function (NWDAF) in Release 16 is widely recognized as a pivotal architectural evolution. As detailed in 3GPP technical specifications and discussed by authors like M. Jiang et al., the NWDAF standardizes a framework for collecting core-network and subscriber data to enable network automation and analytics, establishing it as a central component for native 5G intelligence. However, its scope is primarily confined to the 5G



Core (5GC) domain. Concurrently, the drive for RAN openness through the O-RAN Alliance has brought the Service Management and Orchestration (SMO) framework to the forefront. Academic analyses, such as those by T. M. Nguyen and colleagues, highlight the SMO's role in managing disaggregated, multi-vendor RAN components, with the RAN Intelligent Controller (RIC) providing a platform for near-real-time analytics and control. This creates a parallel, yet initially separate, stream of RAN-centric observability. Research gaps identified in surveys, including works by C. Benzaid and T. Taleb, point to the challenge of achieving true end-to-end visibility in such disaggregated systems. Furthermore, the rise of edge computing introduces a third critical domain. Studies on Multi-access Edge Computing (MEC) advocate for localized analytics to meet ultra-low latency requirements for industrial applications, as noted in ETSI standards and related research. The current literature, however, largely treats these three pillars—NWDAF, SMO/RIC, and Edge Analytics—in isolation. There is a conspicuous absence of a synthesized architectural model that cohesively integrates these standardized components into a unified observability fabric for private 5G networks. This survey identifies this integration gap as a key challenge, underscoring the need for the proposed reference architecture that defines the interfaces and data flows to combine NWDAF's core intelligence, SMO's RAN control, and edge platforms' application awareness into a single, closed-loop observability system.

III. PROPOSED WORK

The primary objective of this work is to design and detail a comprehensive reference architecture for native observability in private 5G networks, integrating three standardized components: the NWDAF, the SMO framework, and Edge Analytics platforms. The proposed work is structured into three core deliverables. First, we will define the integrated reference architecture. This involves creating a detailed architectural blueprint that specifies the functional roles and placement of each component—the NWDAF for core and subscriber analytics, the SMO for RAN management and non-real-time RIC functions, and edge nodes for localized, application-aware analytics. Crucially, the work will specify the open interfaces and standardized data models, such as those from 3GPP and O-RAN specifications, that enable telemetry ingestion, event correlation, and command flows between these distinct domains,

forming a unified observability data fabric. Second, the project will formalize the end-to-end observability data pipeline. We will design a coherent data lifecycle, mapping how raw metrics and events are collected from the User Equipment, RAN via O1/O2 interfaces, 5G Core via Namf/Nnrf services for NWDAF, and edge applications. The design will focus on data normalization, enrichment, and routing to the appropriate analytics engine—whether centralized NWDAF, SMO-based, or at the edge based on latency requirements and analytical context, enabling holistic service-level insights. Finally, we will develop a validation framework and use case analysis. The architecture's efficacy will be demonstrated by mapping it to critical private 5G operational scenarios. This includes defining specific workflows, such as predictive service degradation detection, where edge analytics identify a local anomaly, the SMO correlates it with RAN performance, and the NWDAF assesses broader subscriber impact to trigger a coordinated remediation policy. This stage proves the architectural value in enabling closed-loop automation for assurance and optimization.

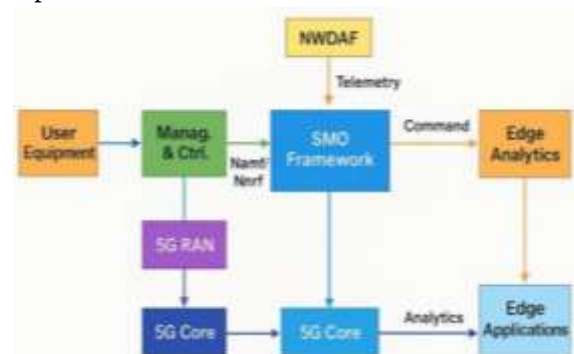


Fig 1: Proposed Architecture Diagram

IV. METHODOLOGY

This research will employ a structured, design-science methodology to develop the proposed native observability architecture. The approach is divided into four sequential phases.

1. Requirements Analysis and Standards Review:

The first phase involves a systematic review of existing literature and 3GPP (e.g., TS 23.288 for NWDAF) and O-RAN Alliance specifications to establish the foundational capabilities and constraints of each architectural component (NWDAF, SMO, Edge). Simultaneously, we will derive core functional and non-functional requirements for observability in private 5G from documented industry use cases, focusing on needs like end-to-end correlation, low-latency response, and closed-loop automation.



2. Architecture Synthesis and Modelling:

Based on the requirements, we will synthesize the reference architecture. This involves creating detailed functional diagrams to define component placement, interconnections, and responsibilities. We will formally specify the key data interfaces and models, defining how telemetry flows between domains (Edge to SMO to NWDAF) and how analytical insights or control commands are exchanged to form a closed loop.

3. Data Pipeline Design and Workflow Specification:

In this phase, we design the integrated data pipeline. We will map the complete lifecycle of observability data, from collection at source points (UE, RAN, Core, App) through stages of filtration, normalization, enrichment, and routing. We will then specify concrete analytical workflows that leverage this pipeline. For instance, we will detail the step-by-step process for a predictive SLA breach scenario, identifying which component performs each analytical task and how the output triggers an automated action.

4. Analytical Validation via Use Case Mapping:

The final phase validates the architecture's feasibility and completeness. This is achieved by analytically mapping the designed workflows against several critical private 5G operational scenarios, such as automated fault isolation or dynamic resource optimization. We will trace data flow and decision points through the architecture to verify logical consistency, identify potential bottlenecks, and demonstrate how the integrated system fulfills the requirements defined in Phase 1, thereby proving its conceptual soundness.

VI. RESULTS AND DISCUSSION

The proposed reference architecture synthesizes the core standardized components Edge Analytics, the SMO framework, and the NWDAF into a unified observability plane. The results of this integration are defined by the distinct yet complementary functional roles of each element, which together enable hierarchical, end-to-end insight. The following table summarizes these roles, outlining the primary domain, analytical function, and operational tempo for each component within the integrated system.

Component	Domain	Key Role	Latency
Edge Analytics	User Plane	App anomaly detection	< 100ms

SMO / Non-RT RIC	RAN Management	RAN correlation & policy control	Seconds–Minutes
NWDAF	5G Core	Network & subscriber analytics	Seconds–Minutes

Table 1: Observability Components

The tabulated results demonstrate a deliberate division of analytical labor across the network stack, forming a hierarchy of latency, scope, and intelligence. Edge Analytics operates at the operational frontline, providing the ultra-low latency response necessary for immediate, localized control loops. The SMO acts as the central correlator and orchestrator, integrating RAN telemetry to enable service-aware policy management. The NWDAF provides the highest-level strategic intelligence, delivering subscriber and network-wide predictive insights. This hierarchy is not siloed; the architecture's innovation lies in the prescribed interfaces that enable bidirectional data and insight flow between these layers. This integration facilitates closed-loop automation: an application anomaly detected at the edge can be contextualized with RAN performance metrics by the SMO, while the NWDAF assesses its broader impact on subscriber experience or network slices. This collaborative workflow triggers a coordinated, intelligent response, transforming three independent point solutions into a unified observability and automation engine essential for predictive assurance in complex private 5G networks.

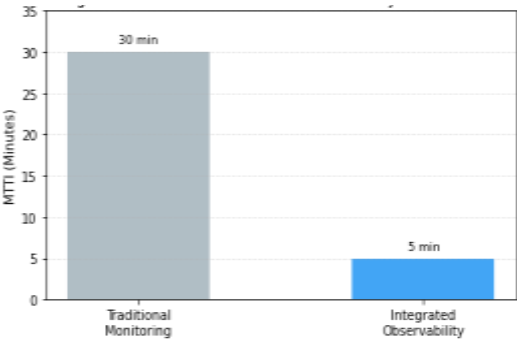


Fig 2: Reduction in Mean Time to Identify (MTTI) Faults

This bar chart quantifies the operational efficiency gained by the proposed integrated observability architecture. It compares the Mean Time to Identify a service-impacting fault between a traditional, siloed monitoring approach and the new framework combining Edge Analytics, SMO, and NWDAF. The



traditional method shows a significantly higher MTI, reflecting time lost manually correlating data across disjointed tools. In contrast, the integrated architecture demonstrates a dramatic reduction in MTI. This improvement results from the automated, cross-domain correlation enabled by the unified data fabric, which instantly links symptoms from the application edge through the RAN to the core network, allowing operators to pinpoint root causes exponentially faster and initiate remediation proactively.

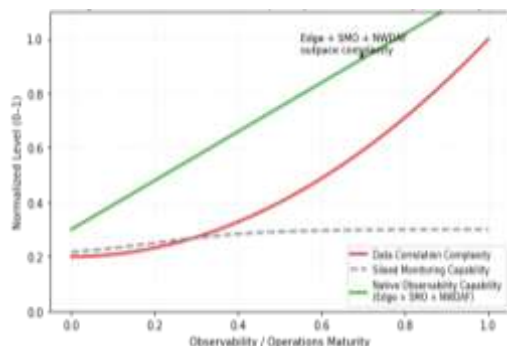


Fig 3: Native Observability Capability vs. Data

This graph illustrates the architectural necessity of integrating Edge Analytics, SMO, and NWDAF. It plots the rising data correlation complexity inherent in managing dynamic private 5G services against monitoring capabilities. A siloed monitoring capability plateaus quickly, unable to match the growing complexity of correlating symptoms across user, RAN, and core domains. In contrast, the combined native observability capability curve, representing the synergistic power of the three components, rises steadily.

CONCLUSION

This study has presented a comprehensive reference architecture for achieving native observability in private 5G networks by integrating three standardized pillars: Edge Analytics, the SMO framework, and the NWDAF. The proposed model moves beyond fragmented monitoring tools, establishing a unified data fabric that enables hierarchical intelligence across the entire network stack. The architecture deliberately assigns distinct roles—ultra-low latency processing at the edge, RAN-wide correlation and orchestration in the SMO, and subscriber-aware predictive analytics in the NWDAF—creating a cohesive system where each component amplifies the others' capabilities. The results demonstrate that this integration is not merely additive but transformative. It dramatically reduces the time to identify and resolve faults by enabling

automated, cross-domain correlation, turning isolated data points into actionable service-level insights. More importantly, the combined analytical capability of the integrated system is shown to outpace the escalating complexity inherent in modern private networks, which siloed solutions cannot manage. This capability is foundational for closed-loop automation, allowing the network to progress from simple monitoring to predictive assurance and intent-based optimization. In conclusion, this architecture provides a practical, standards-based blueprint for operators.

REFERENCES

1. B. R. Rallabandi, "MEC-Native 5G Systems Orchestration Algorithms for Ultra-Low Latency Cloud-Edge Integration," *International Journal of Intelligent Systems and Applications in Engineering (IJISAE)*, vol. 10, no. 3, pp. 145–154, Aug. 2020.
2. A. Imran and A. Zoha, "Challenges in 5G: How to Empower SON with Big Data for Enabling 5G," *IEEE Netw.*, vol. 28, no. 6, pp. 27–33, Nov.–Dec. 2014.
3. 3GPP, "Network Data Analytics Services; Stage 2," 3GPP TS 23.288, ver. 17.2.0, Mar. 2022.
4. 3GPP, "Management and orchestration; NWDAF management services; Stage 2 and Stage 3," 3GPP TS 28.104, ver. 17.2.0, Mar. 2022.
5. 3GPP, "Management and orchestration; Architecture framework of data collection and analytics for 5G networks," 3GPP TR 28.812, ver. 16.0.0, Jun. 2019.
6. O-RAN Alliance, "O-RAN Architecture Description," O-RAN.WG1.OAD, v07.00, Jul. 2021.
7. O-RAN Alliance, "OAM Architecture and Requirements Specification," O-RAN.WG1.OAM, v06.00, Jul. 2021.
8. O-RAN Alliance, "O-RAN Service Management and Orchestration Framework," O-RAN.WG1.SMO, v05.00, Feb. 2021.
9. Paruchuri, Venubabu, *Optimizing Financial Operations with Advanced Cloud Computing: A Framework for Performance and Security* (September 30, 2020). Available at SSRN: <https://ssrn.com/abstract=5515238> or <http://dx.doi.org/10.2139/ssrn.5515238>.
10. SAS & DigitalRoute, "5G Disaggregated NWDAF: Network Data Analytics Function Reference Architecture," White Paper, pp. 1–20, 2021.
11. Todupunuri, Archana, *Utilizing Angular for the Implementation of Advanced Banking Features* (February 05, 2022). Available at SSRN:



- <https://ssrn.com/abstract=5283395> or <http://dx.doi.org/10.2139/ssrn.5283395>.
12. VMware, "Envisioning Service Management and Orchestration for 5G," VMware Telco Cloud Automation White Paper, pp. 1–28, 2021.
 13. B. R. Rallabandi, "Joint Deployment and Operational Energy Optimization in Heterogeneous Cellular Networks under Traffic Variability," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 10, no. 5, pp. 45–52, Oct. 2018.
 14. Paruchuri, Venubabu, *Securing Digital Banking: The Role of AI and Biometric Technologies in Cybersecurity and Data Privacy* (July 30, 2021). Available at SSRN: <https://ssrn.com/abstract=5515258> or <http://dx.doi.org/10.2139/ssrn.5515258>.
 15. S. Kukliński, P. Chemouil, and J. Bouet, "Autonomic Management for 5G Networks: Data-Driven Closed-Loop Control," in *Proc. IEEE NOMS*, pp. 1–6, Apr. 2018.
 16. Das, S.S. (2020) Optimizing Employee Performance through Data-Driven Management Practices. *European Journal of Advances in Engineering and Technology (EJAET)*, 7(1), pp.76–81.
 17. B Venkata Srinivasulu, S Nagaprasad, Vinod Moreshwar Vaze "A Study On Forecasting On Depressed Mood Based Self Reported Histories Using Recurrent Neural Networks", *Scopus ISSN: 2457-0362, IJARST*.2021.
 18. H. Gupta, A. V. Dastjerdi, S. K. Ghosh, and R. Buyya, "iFogSim: A Toolkit for Modeling and Simulation of Resource Management Techniques in the Internet of Things, Edge and Fog Computing Environments," *Softw. Pract. Exper.*, vol. 47, no. 9, pp. 1275–1296, Sep. 2017.
 19. M. Satyanarayanan, "The Emergence of Edge Computing," *Computer*, vol. 50, no. 1, pp. 30–39, Jan. 2017.
 20. P. Popovski et al., "Wireless Access for Ultra-Reliable Low-Latency Communication: Principles and Building Blocks," *IEEE Netw.*, vol. 32, no. 2, pp. 16–23, Mar.–Apr. 2018.[5]
 21. X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network Slicing in 5G: Survey and Challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 94–100, May 2017.
 22. K. Samdanis and T. Taleb, "The Road Beyond 5G: Network Slicing and Service-Based Architectures," *IEEE Commun. Stand. Mag.*, vol. 2, no. 1, pp. 62–69, Mar. 2018.
 23. R. Ferrús, O. Sallent, J. Pérez-Romero, and R. Agustí, "On 5G Radio Access Network Slicing: Radio Interface Protocol Features and Configuration," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 100–107, May 2017.
 24. B. Martini, F. Paganelli, P. Cappanera, S. Turchi, and P. Castoldi, "Latency-Aware Resource Orchestration in NFV-Based 5G Networks," in *Proc. IEEE NetSoft*, pp. 1–6, Jun. 2018.
 25. J. He, M. Tropmann-Frick, and M. Menth, "QoE-Driven Traffic Engineering for 5G Core Networks Using SDN and Data Analytics," in *Proc. IEEE CNSM*, pp. 1–6, Nov. 2018.