

Evolving Challenges and Solutions in Network Management

Jaime Jiménez

jaime.jimenez@ericsson.com

Ericsson, ER

Jorvas, Finland

Scott Mansfield

scott.mansfield@ericsson.com

Ericsson, BNEW TS ST

Jorvas, Finland

Raquel Rodriguez A

raquel.a.rodriguez@ericsson.com

Ericsson, ETAC

Jorvas, Finland

Mikko Pesonen

mikko.pesonen@ericsson.com

Ericsson, ETAC

Jorvas, Finland

Vesa Torvinen

Vesa.Torvinen@ericsson.com

Ericsson, ETAC

Jorvas, Finland

Janne Karvonen

janne.karvonen@ericsson.com

Ericsson, ETAC

Jorvas, Finland

ABSTRACT

This paper presents an analysis of the evolving challenges and emerging solutions in network management. We explore scalability issues in network models and protocols, emphasizing the some limitations of current implementations the need for standardized, extendable information models. Telemetry complexities are discussed, focusing on data quality, diversity, and lineage, and the necessity for efficient data streaming and standardized schemas. Security challenges are addressed in the context of diverse protocols and the transition towards zero-trust architectures, highlighting the importance of unified security mechanisms and continuous updates. We also examine the future of network management, including the integration of generative AI and agentic architectures that adhere to autonomic networking principles, as well as the possibilities of new standard interfaces like CoRECONF. Our findings highlight the imperative for scalable, secure, and interoperable solutions that can adapt to the dynamic demands of modern telecommunications networks.

KEYWORDS

Network Management, Scalability, Telemetry, Security, AI, Zero-Trust, Interoperability

Reference:

Jaime Jiménez, Scott Mansfield, Raquel Rodriguez A, Mikko Pesonen, Vesa Torvinen, and Janne Karvonen. 2024. Evolving Challenges and Solutions in Network Management. In *submissions to the IAB Next Era of Network Management Workshop*, 5 pages.

1 INTRODUCTION

The IAB workshop on the Next Era of Network Management Operations (NEMOPS) serves as a platform for discussion between network operators, protocol experts and the general network management community. This workshop is expected to guide the Internet Engineering Task Force (IETF) standards process. The workshop's primary objectives are to assess past achievements and delineate future requirements for network management operations.

In this paper, we introduce a comprehensive analysis of the current challenges and emerging solutions in network management from the point of view of an SDN controller product. The subsequent sections delve into various aspects of network management and operations. The **Overall Architecture** section 1.1 provides a detailed overview of the standard components within a network management controller, the Ericsson Transport Automation Controller (ETAC). The **Scalability** section 2 examines the challenges of scaling network models and protocols, highlighting the need for standardized models. The **Telemetry** section 3 discusses the complexities of data transmission. The **Security** section 4 raises some security challenges and explains the shift towards zero-trust

This paper is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Copyright (c) 2024 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

IAB NEMOPS WS, December 2024,

architectures. Finally, the **Network Management Evolution** section 5 explores potential future trends, including the role of generative AI and new standard interfaces.

1.1 ETAC overall Architecture

Ericsson Transport Automation Controller (ETAC) [12] is a cloud-native Transport Automation and SDN Controller that leverages artificial intelligence and machine learning to deliver advanced analytics and automation functionalities across microwave, IP, and optical fronthaul networks.

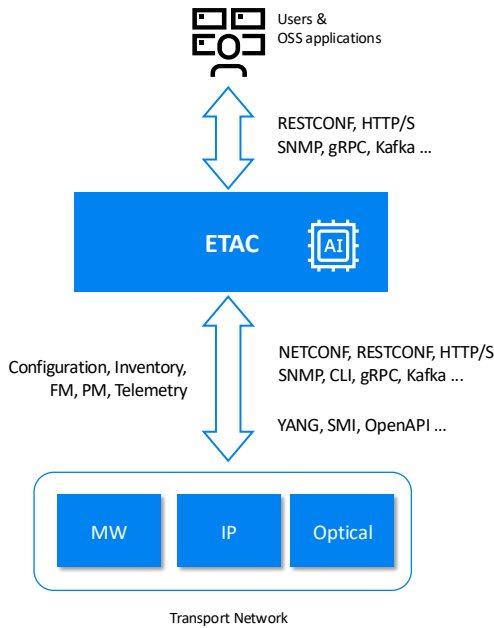


Figure 1: General Architecture

The **Ericsson Transport Automation Controller (ETAC)** [12] adheres to the Open Transport SDN Reference Architecture as delineated by the TIP MUST project, aligning with the principles outlined in the Open Transport Architecture Whitepaper [15]. ETAC supports the roles of both the SDN Domain Controller and the SDN Hierarchical Controller, in accordance with the objectives of the Open Optical & Packet Transport (OOPT) initiative [16]. At high-level, it has two main Integration parts the Northbound Integration part (NBI) and the Southbound Integration part (SBI) both of which use well-known standard interfaces and protocols like **NETCONF/YANG** [3, 11], **RESTCONF** [2], **SNMP** [10], **SFTP** [18] or **HTTP/S** [13] towards APIs.

On the SBI it communicates with the managed nodes and it also is capable of translating between specific device information models and the harmonized, standards-based information model used in a network database. The Network

Intelligence layer builds on top of this harmonized model, implementing the analytics, automation and SDN control application supported in the system.

On the NBI it offers an exposure layer to users over a friendly GUI as well as to various OSS applications, providing access both to the network data and to the functionality supported in the system and enabling integration with other platforms.

ETAC supports both SBI and NBI through implementations that accommodate legacy protocols and information models, whether defined by SOD or specific to vendors. It also provides real-time network observability, facilitating network analytics and closed-loop automation. It currently supports use cases that utilize its built-in real-time observability features to gain network insights through AI/ML and implement closed-loop automation in Transport Networks, all within a zero-trust framework.

2 SCALABILITY

YANG Scalability

Scalability in YANG is challenging, as noted by Boyd [9]. YANG validation can be computationally intensive, and performance degrades with large data stores. Models must be tailored for large-scale devices.

The complexity of YANG is largely driven by its hierarchical architecture. For example, optical equipment can have tens of thousands of interfaces, but current models like YANG often fall short because they use file databases without indexing. There is also concern about the viability of the current implementation of YANG Schema-mount [4] for Optical Network Units (ONU) YANG models, as it seems incompatible with the ONU template technique without editing the YANG modules.

Some modeling challenges can be mitigated by altering the model's structure; however, this often results in backward incompatibility. While YANG features can be adapted or extended to address these issues, doing so introduces additional complexity and also leads to backward incompatibility.

Efficient Data Streaming for Analytics

Legacy network elements predominantly rely on periodic data harvesting, which imposes unnecessary load on the network elements and delays data accessibility. To address this, data sources should implement active streaming of data to post-processing systems immediately upon production, this will also ensure that closed-loop automation systems have access to data in near real-time.

3 TELEMETRY

Quality of data

The alignment of analytics data schemas and metadata is essential for efficient post-processing and analytics. Minor differences, such as variations in timestamp formats, can pose challenges. The adoption of standardized, extendable schemas and encodings for various analytics data types would facilitate streamlined data processing across different vendor systems.

Legacy network elements produce data that may not be fully compatible with modern IT-style post-processing analytics systems. To enhance compatibility, it is beneficial for this data to include embedded metadata providing details such as data originator.

Lastly, while the reliability of telemetry data may be less critical for visualization and manual inspections, it becomes essential when utilized for AI training and assisted network configuration.

Diversity of data

Analytic data is often accessed through a variety of interfaces and standards. This variety presents challenges in streamlining data collection and ingestion processes. This is also true for including metrics which come in various forms like events, alarms, notifications, publications webhooks, logs, and other resource state data.

There is therefore need for an efficient, scalable, and universal transfer system to stream analytics data from its source to post-processing systems.

Lineage of data

It is essential to establish the lineage and integrity of any record used for purposes beyond visualization. Utilizing data with uncertain integrity can introduce new attack vectors, such as enabling adversaries to exploit closed-loop automation by manipulating the input data.

Ensuring data integrity across multiple domains is crucial, especially when analytics data traverses cloud provider PaaS functions before reaching post-processing systems.

State-Data Handling

In passive data sources where analytics data is just exposed (as opposed to streamed) the data has to be actively harvested by external entities, thus there must be APIs or interfaces for collection of such data. In systems where data is streamed, the data sources (often involved in critical functions such as traffic handling) do not need to expose interfaces for harvesting analytics data, which makes the attack surface smaller and system more secure in general.

Modern network elements have a lot of state-data which can not be efficiently leveraged through config model notifications. There are components and elements in the applications/data sources, which one does not to expose in a model,

but still the analytics data about the state of such components should be available for post-processing systems. For example, one ideally would like to keep the model of cloud service realization agnostic, and use the same model for the cloud service, whether the application-instances of those services run in virtual machines or in container. Having to model compute resources in order to convey the state data would make it very tricky to separate realization and model from each other, and keep the model backwards compatible as the application realizations evolve and technologies change.

4 SECURITY CHALLENGES

Security configuration in network management is complex due to the absence of a unified infrastructure. This complexity arises from the need to support multiple security protocols across diverse devices and vendors. Albeit rare, some challenges include:

- **Diverse Protocols:** Network management applications must accommodate various security protocols, such as TLS/DTLS, SSH, and username/password mechanisms.
- **TLS/DTLS:** TLS/DTLS, particularly with client-server certificates, is a promising candidate for unified security. However, some vendors opt for alternative security techniques, such as VPNs.
- **SSH Limitations:** There is no security infrastructure available that would distribute and help verifying the SSH public keys and facilitate flexible re-keying. The security configuration of SSH itself remains to be manual.
- **Username/Password:** Commonly used for basic access and protocols like SNMPv3. Centralized authentication exists but cannot fully replace local authentication due to potential unavailability. Manual configuration often results in poor security practices, as key renewals and password updates are prone to errors and costly.

Legacy security

The transition from standards to widespread network deployment is often very slow, particularly when new standards are to replace existing components. Additionally, management applications often necessitate market adoption or commitment to the new mechanism before its implementation is considered worthwhile. Consequently, phasing out legacy security mechanisms can be challenging.

Zero-Trust Architecture

Zero trust security, or zero trust architecture (ZTA), mandates "never trust, always verify," ensuring no default trust for users or devices, even on permissioned networks. Ericsson develops products that are configured to use secure protocols and configurations by default.

Some operational networks still rely on outdated devices, and any attempt to upgrade or modify these systems is often perceived as a risk, resulting in the neglect of essential software and security updates. While this approach might have been viable within a walled-garden security paradigm, it is incompatible with the zero-trust framework.

5 NETWORK MANAGEMENT EVOLUTION

The evolution of network management could be influenced by advancements in AI and machine learning, as well as potential retrofitting of other protocols intended for constrained environments. These developments promise enhanced automation and scalability for telecommunications networks.

AI Agents

The advent of Large Language Models (LLMs) presents an opportunity for autonomous networking and network management. LLMs are already employed to enhance operator insights, streamline incident management, and generate task-specific code from natural language queries. As the industry advances along the GenAI path, these capabilities will be further augmented by integrating agentic architectures into network AI systems.

LLMs enable the creation of AI agents that are capable of decomposing user intents into executable individual steps and can interact autonomously with other systems through well-known interfaces (e.g., Network and Management APIs). This integration aligns with the autonomic networking principles outlined in RFC7575 [1], facilitating more efficient and adaptive network management.

"The fundamental goal is self-management, including self-configuration, self-optimization, self-healing, and self-protection."

Furthermore, agents can federate and specialize in multi-agent systems, which improve on challenges such as hallucinations, specialization, and scalability. The interest in multi-agent systems was reflected during the past IETF 121 side meeting, as detailed in the ai4network agenda [14].

Agents are particularly useful in the telecommunications sector, where the complexity of specifications and codebases demands innovative solutions. Moreover, the current trend in 5G Networks is marked by a shift towards exposing network functionalities through APIs. This trend facilitates the integration of agentic systems that can dynamically interact with these APIs.

Challenges in this domain include aligning intents from various sources that may have conflicting objectives and ensuring AI outputs are free from hallucinations. The latter is addressed by incorporating fact-checking, dataset tagging,

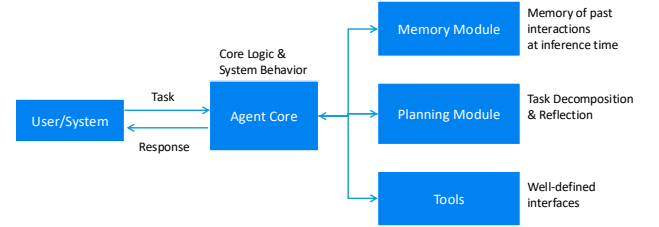


Figure 2: Agent Architecture

human-in-the-loop solutions, and other verification mechanisms. On top of that it is likely that there will be incremental improvements on each of the components of an agent, both in planning, memory and tooling.

Moreover, there are limitations in terms of control of the most advanced LLMs which are the foundation of these types of agents, it is likely that privacy concerns and lack of in-house LLM deployments limits the usage.

CoRECONF

There is potential for retrofitting Internet of Things (IoT) oriented protocols on telemetry or management-type signaling within the network management domain. For example in UDP environments and in environments where compression and smaller payloads are welcomed.

In particular, CoRECONF is utilized by constrained devices in Low-Power and Lossy Networks, which are typically composed of numerous embedded devices with limited power and memory.

The the Constrained Application Protocol (CoAP) [17], a component of CoRECONF, is primarily utilized in IoT applications, it has also been specified for signaling DDoS-related telemetry [8] and [7], as documented by the now-concluded DOTS Working Group.

Within the network management community, the -CONF ecosystem is predominantly characterized by the widespread deployment of NETCONF and RESTCONF for network management tasks. In contrast, CoRECONF is not as widely know. The main differences being the application protocol and the serializations:

- **NETCONF** [11]: Serializing YANG over a stateful TCP connection.
- **RESTCONF** [2]: Serializing YANG over stateless HTTP.
- **CoRECONF** [6]: Serializing YANG modules in a CBOR [5] map over stateless CoAP.

6 CONCLUSIONS

This paper has explored the evolving landscape of network management, highlighting the challenges and solutions in scalability, telemetry, and security. The integration of advanced technologies such as AI and machine learning within platforms like the Ericsson Transport Automation Controller (ETAC) demonstrates the potential for enhanced analytics and automation in network operations. However, the complexity of legacy systems and the diversity of data sources present significant hurdles that require innovative approaches, such as standardized schemas and efficient data streaming mechanisms.

The transition towards zero-trust architectures and the adoption of generative AI indicate a shift towards more secure and autonomous network management systems. These advancements necessitate a collaborative effort between industry stakeholders and standardization bodies to ensure seamless integration and widespread adoption. As the network management domain continues to evolve, the focus must remain on developing scalable, secure, and interoperable solutions that can adapt to the dynamic demands of modern telecommunications networks.

7 ACKNOWLEDGMENTS

We would like to express our gratitude to Ericsson for their support of this work. Special thanks to Jari Arkko for his insightful review.

REFERENCES

- [1] M. Behringer, M. Pritikin, S. Bjarnason, A. Clemm, B. Carpenter, S. Jiang, and L. Ciavaglia. 2015. *Autonomic Networking: Definitions and Design Goals*. Request for Comments RFC 7575. RFC Editor. <https://www.rfc-editor.org/rfc/rfc7575> Informational.
- [2] A. Bierman, M. Björklund, and K. Watsen. 2017. *RESTCONF Protocol*. Request for Comments RFC 8040. RFC Editor. <https://www.rfc-editor.org/info/rfc8040> Standards Track.
- [3] M. Björklund. 2010. *YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)*. Request for Comments RFC 6020. RFC Editor. <https://www.rfc-editor.org/rfc/rfc6020> Standards Track.
- [4] M. Björklund and L. Lhotka. 2019. *YANG Schema Mount*. Request for Comments RFC 8528. RFC Editor. <https://www.rfc-editor.org/info/rfc8528> Standards Track.
- [5] C. Bormann and P. Hoffman. 2022. *Encoding of Data Modeled with YANG in the Concise Binary Object Representation (CBOR)*. Request for Comments RFC 9254. RFC Editor. <https://www.rfc-editor.org/info/rfc9254> Standards Track.
- [6] C. Bormann, P. van der Stok, and A. Sehgal. 2023. *CoAP Management Interface (CoMI)*. Internet-Draft draft-ietf-core-comi. <https://datatracker.ietf.org/doc/draft-ietf-core-comi/> Work in Progress.
- [7] M. Boucadair and J. Shallow. 2023. *Distributed Denial-of-Service Open Threat Signaling (DOTS) Signal Channel Configuration Attributes for Robust Block Transmission*. Request for Comments RFC 9362. RFC Editor. <https://www.rfc-editor.org/info/rfc9362> Standards Track.
- [8] M. Boucadair, J. Shallow, and T. Reddy.K. 2021. *Distributed Denial-of-Service Open Threat Signaling (DOTS) Signal Channel Specification*. Request for Comments RFC 9132. RFC Editor. <https://www.rfc-editor.org/info/rfc9132> Standards Track, Obsoletes RFC 8782.
- [9] A. Boyd. 2023. Scalable YANG. IEEE 802.1 YANGsters. <https://www.ieee802.org/1/files/public/docs2023/yangsters-boyd-scalable-yang-1123-v01.pdf> Presentation on scalable YANG models.
- [10] J. Case, M. Fedor, M. Schoffstall, and J. Davin. 1990. *A Simple Network Management Protocol (SNMP)*. Request for Comments RFC 1157. RFC Editor. <https://www.rfc-editor.org/rfc/rfc1157> Standards Track.
- [11] R. Enns, M. Björklund, J. Schönwälder, and A. Bierman. 2011. *Network Configuration Protocol (NETCONF)*. Request for Comments RFC 6241. RFC Editor. <https://www.rfc-editor.org/rfc/rfc6241> Standards Track.
- [12] Ericsson. 2024. Ericsson Transport Automation Controller (ETAC). <https://www.ericsson.com/en/portfolio/networks/ericsson-radio-system/mobile-transport/ericsson-transport-automation-controller>. Accessed: 2024-11-07.
- [13] R. Fielding and J. Reschke. 2014. *Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing*. Request for Comments RFC 7230. RFC Editor. <https://www.rfc-editor.org/rfc/rfc7230> Standards Track.
- [14] Daniel King and Weiqiang Cheng. 2023. ai4network: An IETF Side Meeting for the Discussion of AI and its Applicability to the Network. IETF 121 Side Meeting. <https://github.com/danielkinguk/ai4network/blob/main/ietf121/agenda-121.md>
- [15] Telecom Infra Project. 2023. Open Transport Architecture Whitepaper. Whitepaper. https://cdn.mediaialet.com/usva/telecominfraproject/03V-53HVHE2_sr3_nk47_Q/WP46tLiuS0CDkcG5S6Etug/Original/OpenTransportArchitecture-Whitepaper_TIP_Final.pdf Accessed: 2023-10-15.
- [16] Telecom Infra Project. 2024. Open Optical & Packet Transport (OOPT). Telecom Infra Project. <https://telecominfraproject.com/oopt/> The Open Optical & Packet Transport Project Group aims to accelerate innovation in optical and IP networks, enhancing connectivity for global communities..
- [17] Z. Shelby, K. Hartke, and C. Bormann. 2014. *The Constrained Application Protocol (CoAP)*. Request for Comments RFC 7252. RFC Editor. <https://www.rfc-editor.org/info/rfc7252> Standards Track.
- [18] T. Ylönen and C. Lonvick. 2006. *The Secure Shell (SSH) Transport Layer Protocol*. Request for Comments RFC 4253. RFC Editor. <https://www.rfc-editor.org/rfc/rfc4253> Standards Track.