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 a comprehensive analysis of the evolving challenges and solutions in network management, focusing on scalability, telemetry, security and possible future technological additions. It uses the Ericsson Transport Automation Controller (ETAC) as a typical deployment of an SDN controller and Transport Automation system. The paper highlights the complexities of legacy systems and the need for standardized models and efficient data streaming. It also explores the transition towards zero-trust architectures. Lastly, we examine potential additions to the network management domain, specifically the integration of generative AI and agentic architectures that adhere to autonomic networking principles, as well as the possibilities of retrofitting modern constrained protocols in the networking domain.

KEYWORDS

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

IAB NEMOPS WS, December 2024,
.

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 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.

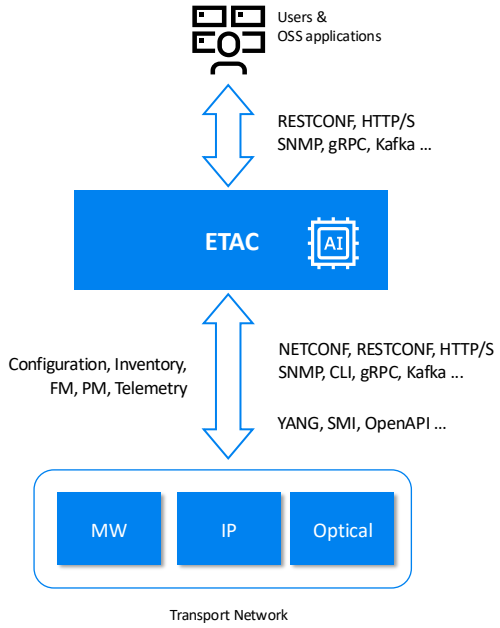

Figure 1: General Architecture

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 Northound 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 a Standards Developing Organization (SDO) or by 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

Protocol efficiency

In high-rate telemetry use cases, line protocol efficiency is a key consideration. With NETCONF, the verbose nature of the XML encoding implies both added processing overhead

in server and client, and increased bandwidth requirement on the DCN channel. Techniques for reducing the processing and bandwidth requirements should be considered in future work. One example of the importance of efficiency is is gRPC, which has gained popularity in telemetry applications largely due to its line protocol efficiency.

Similarly, the choice of transport protocol has a key impact on the processing and bandwidth requirements. TCP, even though more reliable, is heavier in processing and on the line so often UDP would be preferred for transport in telemetry applications. With UDP, however, additional mechanisms may be needed on application layer to compensate for the unreliable transport and to handle e.g. packet drops and reorderings, in order not to lose critical data, and these mechanism may in turn result in losing the benefits of the lighter transport layer. There's thus a fine balance between when to use UDP vs. TCP and the choice should be based on careful consideration of all these factors.

Quality and integrity of the data

The alignment of telemetry data across different sources and subsystems is essential for efficient post-processing of the data. This calls for more coordination in the development of YANG data models, e.g. in order to be able to represent the necessary associations between YANG modules developed in different groups and the schemas of the records, so that the data can be linked correctly and efficiently in post-processing. Also what should be taken into account is the lineage and integrity in the streamed data records, e.g. to not introduce new attack vectors by enabling adversaries to exploit closed-loop automation by manipulating the telemetry data.

Flexibility

The amount of configuration, state and performance data available and produced in modern network elements is ever-increasing, and ability to have all this data available for network management, analytics and automation applications is crucial in managing the complexity of networks today and in the future. Thus, the protocols and mechanisms developed for streaming telemetry should be optimized to support high-rate streaming of large volumes of data efficiently.

The potentially high volume of data also calls for ability to have fine-grain control on what data gets streamed. Different applications are interested in different subsets of the data, may need it at different frequency or their needs may e.g. be adaptive according to the condition of the network. Also, there's often a tradeoff between the needs of an application and what can be delivered e.g. due to the limited DCN bandwidth available in the network. Being able to exactly choose what gets streamed and when helps to overcome these challenges.

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:** Security infrastructure to distribute and help verifying the SSH public keys and to facilitate flexible re-keying is not commonly in use, probably due to complexity.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.

In legacy networks, sometimes the transport network is assumed to be secured, trusted, so that lower security can be accepted. This may work in some networks but the trend is definitely towards zero-trust. Regulators and local legislation is also setting new requirements for critical infrastructure, such as telecom networks. The interest is not only on the security capabilities required from the network devices but also development processes, documentation, FOSS usage, security assurance and various aspects of the way how networks are deployed and managed.

5 NETWORK MANAGEMENT EVOLUTION

The evolution of network management and automation 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/ML/Analytics for enhanced observability and network insights

Legacy network management platforms ingest different types of PM, FM, events data generated by network elements and in most cases legacy network managers leave the interpretation and correlation of that data to humans.

AI/ML techniques like anomaly detection, classification, prediction, etc can be run over time series data, network topology and inventory data to detect and predict network events by looking at anomalies, patterns in data, forecasting trend, etc.

The AI/MLs/Analytics generated data can then be used to produce network insights in general, being some of those also actionable (closed loop automations).

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

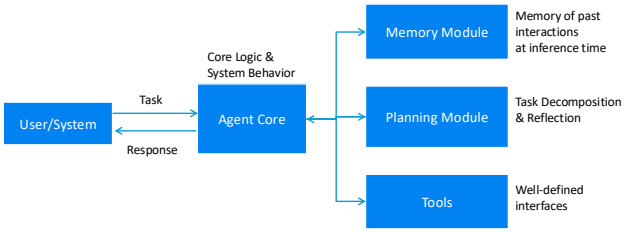

Figure 2: Agent Architecture

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.

The multi-agent domain could also have applications in networks, CSP Transport Networks are typically very heterogeneous with a combination of Transport domains (Microwave, IP, Optical) and equipment vendor. Therefore domain/vendor specific agents can be trained. SDN platforms can then perform orchestration the specialized agents to achieve overall better performance in heterogenous networks.

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, 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 evinronments 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 known. 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.

RFC9254's encoding approach is particularly advantageous in scenarios where minimizing data size is crucial, as it efficiently maps YANG data structures into CBOR maps, greatly compressing the contents of configuration datastores, state data, RPC inputs and outputs or event notifications.

6 CONCLUSIONS

In conclusion, the evolving network management landscape offers both challenges and opportunities. ETAC exemplifies an automation controller utilizing IETF's protocol suite, demonstrating how the industry can tackle the complexities of legacy systems. YANG's scalability is limited by its hierarchical complexity and lack of indexing, resulting in performance issues and backward incompatibility when models are adapted for large-scale devices. The zero-trust architecture necessitates a unified approach to security protocols and configurations to mitigate vulnerabilities in legacy systems. Finally, AI/ML techniques enhance network management by automating data interpretation and generating actionable insights. LLM-agents may facilitate autonomous networking and execute user intents, while CoRECONF, leveraging CoAP and CBOR, provides efficient data compression and reduced payloads that could be leveraged in the general networking domain

7 ACKNOWLEDGMENTS

The authors would like to express their gratitude to Ericsson for supporting this work. Special thanks to Jari Arkko for his review and insightful feedback.

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.mediavalet.com/usva/telecominfraproject/03V-53HVHE2_sr3_nk47_Q/WPd6tLiuS0CDkcG5S6Etug/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.