A Model-based Approach to Multi-domain Monitoring Data Aggregation

Antonio Pastor[1,2], Diego R. López[1,*], Jose Ordonez-Lucena[1], Sonia Fernández[1] and Jesús Folgueira[1]

[1]*Telefónica I+D – Madrid, Spain*
[2]*Universidad Politécnica de Madrid – Madrid, Spain*
E-mail: antonio.pastorperales@telefonica.com; diego.r.lopez@telefonica.com; joseantonio.ordonezlucena@telefonica.com; fernandez.sonia@outlook.es; jesus.folgueira@telefonica.com
[*]*Corresponding Author*

Received 23 November 2020; Accepted 06 March 2021;
Publication 03 June 2021

Abstract

The essential propellant for any closed-loop management mechanism is data related to the managed entity. While this is a general evidence, it becomes even more true when dealing with advanced closed-loop systems like the ones supported by Artificial Intelligence (AI), as they require a trustworthy, up-to-date and steady flow of state data to be applicable. Modern network infrastructures provide a vast amount of disparate data sources, especially in the multi-domain scenarios considered by the ETSI Industry Specification Group (ISG) Zero Touch Network and Service Management (ZSM) framework, and proper mechanisms for data aggregation, pre-processing and normalization are required to make possible AI-enabled closed-loop management. So far, solutions proposed for these data aggregation tasks have been specific to concrete data sources and consumers, following ad-hoc approaches unsuitable to address the vast heterogeneity of data sources and potential data consumers. This paper presents a model-based approach to a data aggregator framework, relying on standardized data models and telemetry protocols, and

Journal of ICT Standardization, Vol. 9_2, 291–310. *River Publishers*
doi: 10.13052/jicts2245-800X.9210

integrated with an open-source network orchestration stack to support their incorporation within network service lifecycles.

Keywords: Data, source, consumer, framework, aggregation, closed-loop, automation, metadata.

1 Introduction

The evolution of networks is making them increasingly complex as shown by, for example, the densification of network-attached devices (e.g. Industry IoT, connected cars, etc.) and the need for heterogeneous access technologies to satisfy user experience demands. This increasing complexity implies management challenges that require a transformation of the current management and operations and management techniques, aiming at zero-touch management.

To achieve zero-touch management, the application of closed-loop control is the most promising one regarding automation goals. Closed-loop control intends to apply a well-established corpus around the discipline of Automatics [1], combined with mechanisms such as AI techniques, and leveraging the additional degrees of freedom for network service and functional management enabled by the recent trends in *software networking*, with Network Function Virtualization (NFV) [2] and Software-Defined Networking (SDN) [3] as flagship technologies.

Closed-loop control has as essential requirement a timely and trustworthy flow of data about the entity being managed. The specific implications of networking regarding heterogeneity, pervasiveness and topology awareness, among others, have made the availability of these data flows challenging, if available at all in an open context.

In this paper, we address the collection stage for closed-loop automation (see Figure 1), developing a solution relying on the ZSM approach for multi-domain data aggregation. For this process collection, based on the transfer of data from sources to consumers for storage and further analysis, we apply a model-based approach to define the metadata specifying sources (e.g. individual NFs, databases, specific devices, cloud and edge infrastructure...) and consumers (e.g. dashboards visualizations, big data processing, real-time analytics, machine learning models...). The ZSM framework use the metadata models not only to collect data from sources and make them available to consumers, but also to perform the necessary transformations to make the data from any source applicable in conjunction with data from other sources.

2 Data Aggregation in Closed-loop Automation

2.1 The Concept of Control Loop

A Control Loop (CL) is a building block for the management of network and services. It is a type of control mechanism that uses feedback to monitor and regulate itself to achieve a specific target. The basic principle of any CL is to adjust the value of a measured or observed variable (expressed as for example an attribute) to equal the value of a desired goal (expressed as for example an attribute).

For the deployment and operation of a CL, the following is required:

- The *managed entity*, which is the target over which the CL is to be applied. In the network environment, this would correspond to a network function or a network service/slice.
- The *producers of the measurements or observations*, which defines the input data based on which the CL acts. The management entity and/or other data sources can be act as producers.
- The goal, which specifies the desired state/behavior that the CL shall maintain over the managed entity, throughout its lifecycle. Once the CL is configured with this goal, the managed entity is adjusted accordingly.

For the CL to act on input in the context of the set goals, the CL follows a set of steps that continuously consume and produce information from each other in a loop that follows this sequence: monitor, analyse, decide and execution. Depending on whether a human operator intervenes inside this loop, a CL can be an open CL (human-assisted CL) or closed CL (no human intervention). In this paper, we focus on closed CLs and the means allowing for their automated deployment and configuration, typically referred to as Closed-Loop Automation (CLA).

2.2 Closed-Loop Automation

Although there is a number of proposals on how a closed CL can be internally articulated, in this paper we consider an OODA (Observe-Orient-Decide-Act) based model for the CLA concept, as defined in [4]. As seen from Figure 1, the use of this model allows defining a closed CL as a composition of four stages: collection, analysis, decision, and execution. *Collection* is responsible for collecting and pre-processing data from the managed entity. Data from external sources that might influence the behavior of the managed entity can also be incorporated in this stage. The *Analysis* stage is responsible for deriving insights from available data from the collection stage, as well as

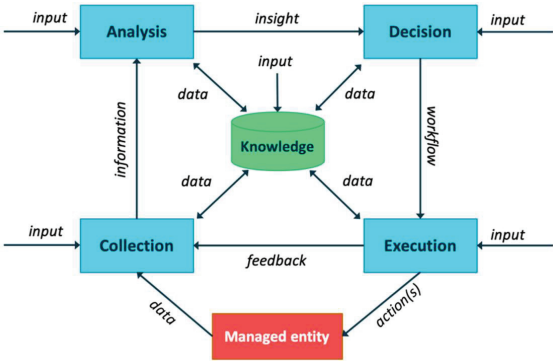

Figure 1 The concept of closed CL.

historical data. This stage provides a diagnosis of the status of the managed entity, determining if it behaves as required (e.g. KPIs are met) and if not, identifying the cause of misbehavior. The *decision* stage is responsible for deriving workflows from insights derived in the analytics stage. The workflows define which actions should be taken over what entities, and in which order, all this in a technology-agnostic manner. Finally, the *execution* stage is responsible for enforcing defined workflows, translating high-level actions into configuration commands or lifecycle management operations (e.g. scaling in/out).

To provide feedback signaling between the four stages, and thus context-awareness to CLA, the *knowledge* artifact is included. Note that this artifact is not a stage itself, but rather a means for storing and retrieving data that is shared between the stages within a CL (as well as between other running CLs). Examples of these data include configuration, operational and historical data.

2.3 Data Aggregation

Following the above-referred model, the CL operates based on the input data (i.e. measurements and observations captured at the collection stage. The producers of these data typically correspond to different sources, whose number and heterogeneity is use case-dependent. In any complex enough system, the number of potential monitoring data sources and their characteristics becomes extremely high and, in the case of entities with many cooperating autonomous components, it creates a *combinatorial explosion* that yields control unfeasible, specially if CLA (i.e. no human operation in

the loop) is pursued. Operator-managed networks constitute a paradigmatic example of these complex systems, especially if a full end-to-end (E2E) control is intended for all the services running on it, including infrastructural and communication services, together with supporting network slices.

In these networks, involved data sources may have different capabilities in terms of:

- Data management. To make data available for consumption, some data sources rely on pull-based methods (e.g. Syslog, IPFIX, SNMP), while there exist others that rely on push-based methods (e.g. streaming telemetry). Unlike polling, which requires explicit requests at continuous polling intervals, push-based solutions allow pushing data off of the device to data consumers in a more efficient fashion, according to a well-defined subscription between both entities. The subscription allows the data consumers (i.e. subscriber) to subscribe to data models, and data sources (i.e. publisher) to push the data to the collector for the subscribed model.
- Data accessibility, with data sources applying a great variety of access control, confidentiality and integrity methods.
- Data availability. In many cases, timing constraints on both the data and their processing have to be considered. The validity of data depends on the time when they are collected and accessed, and the correctness of a process depends on whether it completes on time, and whether the continuity of the data flow can be guaranteed to avoid the *starvation* of the elements in the CL.

According to this rationale, it is clear that data received from operator-managed networks may come in many different formats, from unstructured to strictly based on standard models, being transported according to a great variety of mechanisms, accessed using diverse credentials and be available at different time frames. In these multi-technology, multi-vendor environments, defining ad-hoc solutions for each data source or specific adaptation mechanisms for each data consumer is not acceptable, and would make the application of closed CLs in real network environments highly unfeasible. Indeed, CLA really brings OPEX reduction for operators only if the designed CLA solutions are openly applicable and reusable in different scenarios. In other words, this means that any closed CL suitable for E2E network environments has to evolve to support complex multi-domain scenarios, both at the technology (e.g. wireless and wireline access, optical and packet transport, container- and VM-based cloud environments) and management levels,

including support for data collection and data processing able to integrate loosely coupled data sources and retrieving procedures.

To achieve the above, *data aggregation* mechanisms are an essential component, bridging the collection and analysis stages in the CL model described in Figure 1, thus providing a consistent and manageable set of information for further processing in the decision stage. In this work, a data aggregation framework based on the definition of metadata for both data sources and data consumers is proposed. By allowing for an open, model-based integration of sources and consumers of different nature, this framework may reduce the complexity of incorporating new data sources and consumers stages, thereby increasing the applicability of closed CLs to E2E network and service management, while reducing the complexity of incorporating new data sources and data consumer stages.

3 Applicable Standards

Model-driven data management is based on the idea of applying modelling languages to formally describe data sources, defining their semantics, syntax, structure and constraints on the objects they are associated to. Model-driven data management considers data models, transport protocols and data coding languages as independent layers, easing the aggregation of any new protocol and/or coding that follows the same principles of model-driven data management.

As the reference data model language for network management, YANG [5] has been selected as the language for data modelling. YANG allows to create a data model, define data organization in that model, as well as its constraints. Although model-driven date management started with the NETCONF [6] protocol and the XML [7] coding, there are different solutions that make possible their implementation.

Figure 2 depicts the different components that are available for implementing model-driven data management solutions. Once the YANG data model is defined and implemented, a client (network management system – NMS, application, network orchestrator, network controller, etc.) can select the most appropriate encoding (XML, JSON [8], Protobufs [9], etc.) and a the most appropriate transport protocol (NETCONF, RESTCONF [10] or gRPC/gNMI [11]).

From an architecture point of view, both client and server are driven by the content of the YANG modules where models (and constraints on data) are defined. As a result of parsing the modules, the client is aware of the

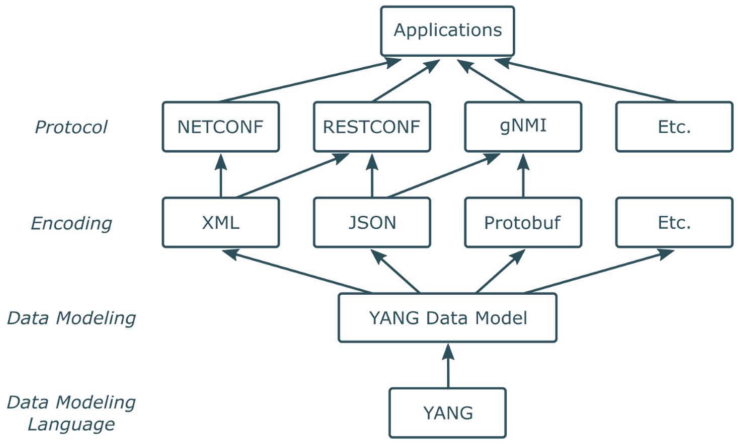

Figure 2 Model-driven data management components.

data the server (data source, device) can provide, and the server knows which are the rules and behavior it should abide by when providing them. The server includes the definition of the modules as metadata, available to the protocol engine that implements the chosen network management transport protocol, and that processes the incoming requests from the clients. Apart from processing these incoming requests, this engine uses the metadata to parse and verify any request, perform the requested operation, and return the result to the client.

As described above, YANG models are commonly used to describe data that is provided by data sources (the servers) but, since the data aggregation framework described in this paper acts as a server for the data consumers attached to it, YANG models are equally used to define the models these consumers will use to access the data mediated by the framework.

The ETSI ISG CIM (cross-cutting Context Information Management [12]) approach is based on the *context* concept. Context information is considered to be any relevant information about entities, their properties, and their relationships with other entities. Entities may be representations of real-world objects but may also be more abstract notions such as a legal entity, a network function or a group of other entities. Relationships between entities can be modelled as having specific properties, and the whole set of entities and relationships are modeled according to a labelled property graph, providing a theoretical foundation for automated reasoning about the characteristics of the systems they represent. Context information is exchanged amongst applications, context producers, and context brokers.

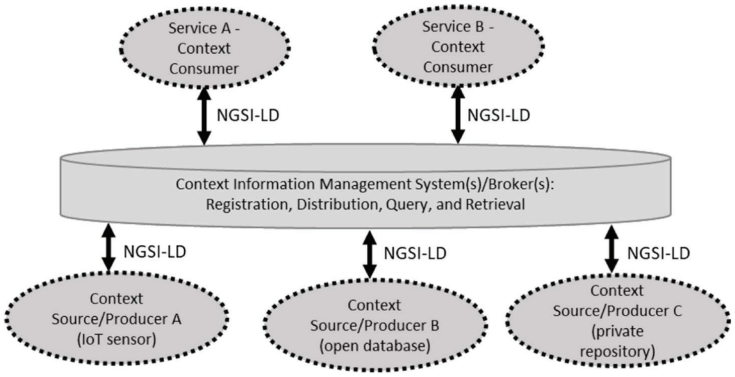

Figure 3 Interconnection of context information (Source: [15]).

CIM considers the exchange of data and metadata across systems a crucial enabler for different applications, to allow them to better collect information from different origins, filter information from "data lakes" and to create derivative information or decisions. CIM considers provisioning of provenance, data quality, access control and other features as part of its NGSI-LD protocol [13]. The acronym NGSI stands for "Next Generation Sensors Initiative", as context information was initially focused on collecting sensor data. The LD (for "linked data") refers to the extension that includes links to metadata references. The NGSI-LD API is based on linking available data within property graphs that reference data definitions (ontologies), such as those in SAREF [14].

The application of the CIM framework has been focused on IoT applications so far, but given its support for binding context (data) sources and consumers, we are incorporating network monitoring data as another class of context to be addressed within CIM, relying primarily on YANG models, though other modeling mechanisms (SNMP MIBs [16], time series databases like Prometheus [17], IPFIX data flows [18]...) are being considered as well.

To take full advantage of the data-enabled approach to network management in the automation of next-generation networks and service management, network service providers are required to transform their current operation support systems (OSS), adopting more flexible architectural approaches that allows addressing the integration challenges and scalability burdens that the execution of CL control mechanisms will bring. A key step in this transformation path is the use of service-based architectures (SBA) [19], based on replacing traditional interface reference points with functionality delivered as

a service, usually requested through a Web Service interface. Unlike traditional telecom-style approaches, this new architecture style places emphasis on the services provided by individual architectural components rather than on the relationships between pre-defined pairs of architectural components. This paradigm shift, when applied at the management layer, allows for structuring the whole OSS as a set of components providing well-defined APIs suitable to be used by other components, or to be invoked by external elements, performing an *API-fication* of the entire OSS with minimal integration, which results in much more scalability support and room for further service innovation.

The ETSI ZSM framework [20] is an architecture solution that builds upon the SBA principles. In this service-based management architecture, the operation is structured around management domains (MDs) that communicate and interact with each other by means of an integration fabric. A MD is formed of a federated set of management services (e.g. orchestration services, control services, analytics services, intelligence services, etc.) that are jointly used for the deployment and operation of instances from a given managed entity. For a network service provider, the ZSM framework considers a set of *network MDs*, operating network functions/services from individual network domains (e.g. access network MD, core network MD, transport network MD, cloud network MD) and one *service MD*, providing an overarching view of E2E services, including cross-domain network slices. To fully automate the operation of their managed entities, each MD, can create one or more closed loops for different tasks like self-configuration, self-healing or self-optimization, using its internal mechanisms and all the services from other domains it is able to consume. It is worth noting that these closed loops can be created as well by the E2E service MD, thus providing full multi-domain E2E automation.

The data aggregation framework is following a SBA approach for the definition of data sources (incorporating the necessary *source agents* whenever required) and data consumers, and it is committed to fully support the ZSM approach, with special emphasis on services characterized as Intelligence, Analytics, Data Collection, and the Data Services themselves.

In order to demonstrate and build a proof of the concepts introduced in this paper with regard to multi-domain data-enabled management supported by metadata models, two core open-source components have been combined: the Orion context broker [21], as the base for managing the lifecycle of context information and its availability, and the Open Source MANO (OSM) stack [22], as a network service orchestration platform allowing for the

deployment of the platform elements and the model-based management of data flows. As a practical case, the data aggregation platform built for this initial proof of concept, is being evolved to support the security architecture described in the next section.

4 Applying Data Aggregation to 5G Security

This section describes a representative example of our model-driven data aggregation platform, applied to a security architecture based on the ETSI ISG ZSM framework. We have combined two key areas where network data analytics and aggregation are fundamental components. First the 5G network, considering a holistic view covering all the involved network domains: Access, Transport and Core. And second, one of the most demanding application in data analytics and intelligent response: network security.

The ETSI ISG ZSM framework fits very well with 5G networks, now that 3GPP promoted the adoption of SBA and the use of the related SBI (Service Based interfaces). Both, ZSM and the 5G Core specification share concepts such as message buses and data service discovery. Additionally, different domains in 5G are evolving to more flexible and efficient methods for managing the network, mainly dependent on data collection and aggregation. In the Radio Access Networks (RAN) domain, the optimization of the radio resources is improved with self-organizing network (SON) solutions, that depends on radio devices monitoring and data collection. Also, the Transport domain that covers fronthaul, backhaul and aggregation segments to provide connectivity towards 5G core and ISP data network areas, follows the adoption of a management and control plane based on Software-Defined Networking (SDN). SDN architecture allows straightforward mechanism to collect data based on standardized telemetry interfaces (e.g. NETCONF at the IP plane, or T-API at the optical plane), easing the aggregation for automation based on closed-loop mechanisms. Finally, infrastructure providers (whether for central offices, edge data centers or cloud) are adopting the NFV paradigm for 5G Core, thus applying virtualization and specific management and orchestration (MANO) functions. Relevant data collected from the NFVI (NFV Infrastructure) and VNFs will allow to make analysis and take decision, expressed as network policies, that will be enforced, through MANO functions.

Data collection, aggregation and analysis are techniques already exploited in the security field for a long time, with Open Source Intelligence (OSINT) techniques [23] as a representative example, where data from public sources,

is aggregated to generate insights and detect attacks. Network traffic related information is essential to leverage advanced machine learning techniques in the cybersecurity area [24], where the quality in the data collected and how it is aggregated impacts in the performance of the detection. ETSI ISG CIM concepts, introduced in Section 3, is proposed as the solution to provide the context in the security data information collected from the network, through NGSI-LD API, to help in the aggregation process.

In INSPIRE-5Gplus [25], a general architecture (Figure 4) to provide an end to end security capacity in 5G networks is introduced. The architecture is strongly aligned with the ZSM framework, adopting its key capabilities (Section 3). The support of multiple 5G domains, with local security intelligence at each domain, is made possible by means of integration fabrics, interconnecting different security functions through a transversal E2E Security Management Domain to coordinate security intelligence and enforcement over different domains.

The components involved in the process of data generation, collection and aggregation are underlined in Figure 4, where a simplified version of INSPIRE5G-plus high-level architecture is shown. The *Security Data Collector* service, based on the combination of Orion and OSM described in the previous section, sets up and launches the mechanisms for collecting and

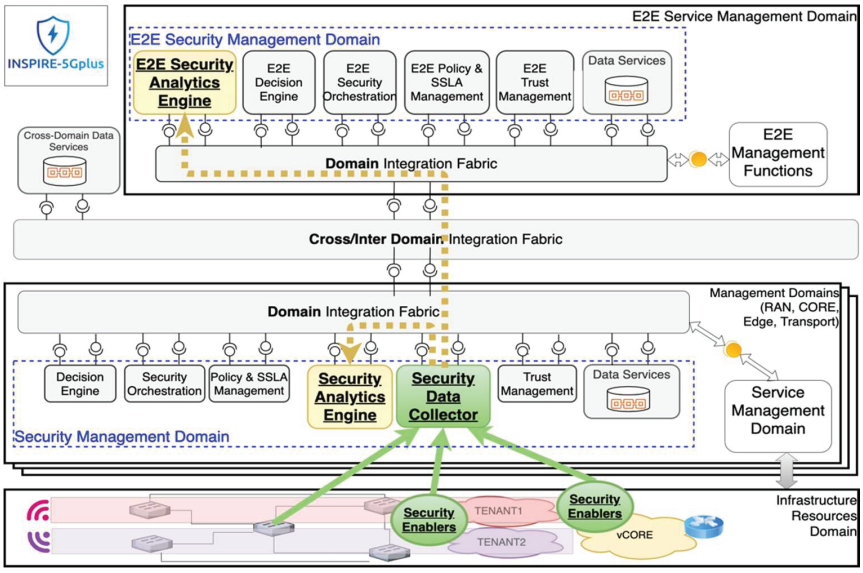

Figure 4 Data aggregation in INSPIRE-5Gplus architecture (simplified).

aggregating data generated from the different network resources (e.g. *security enablers or* network devices) in the *Infrastructure Resources Domain (solid arrows)*. These mechanisms can rely on different message exchange models: publish/subscribe, point-to-point, request-response, etc., and different types of protocols (see Figure 2). The natural variety of data sources (e.g. network device telemetry, processes logs, VNF security events, Firewalls logs, etc.) requires the capacities of the data aggregation framework described in previous Section 2. The Security Data Collector adopt this role. This way, the data aggregation framework is used as the source of information for advanced functions in higher level, such as the *Security Analytics Engine* service that provide machine learning analytic outputs, or to *Data Services* that provide persistent storage for additional analysis of different security functions in each domain or at the E2E level. The communication process is done over a *common domain integration fabric.*

A specific security problem studied with this framework is how to detect attacks and network misuse in the presence of pervasively encrypted traffic over 5G networks. Control plane traffic uses 5G Service Based Interfaces (SBI), which specifies the use of Transport Layer Security (TLS) as the base protocol implementation for RESTful APIs. Data plane applications also rely mainly upon the TLS protocol. One specific attack in this area is the crypto-mining activity, where victim's IT resources (e.g. the ones in the 5GCore or vertical edge services) can be exhausted [26]. Solutions based on ML models to detect the attack have been proposed [27], but this attack can evolve to support TLS as a channel of communication. Using the INSPIRE5G-plus reference framework it is possible to deploy network probes (as security enablers), in different infrastructure resource domains, to capture network traffic and generate customized metrics, collect and aggregate them through the Security Data Collector. This data is progressed, through the integration fabric, to the Security Analytic Engine to feed an improved ML algorithm. Each security enabler generates data with standard protocols such as NetFlow v9 [28] or IPFIX, but also provide additional contextual data with 12 statistical information of TCP flows [29]: Round Trip Time (RTT), TCP protocol flags (such as SYN or ACK) , windows scale or Maximum Segment Size (MSS) per flow. The aggregation of this supplementary data source has already demonstrated a clear increment in the performance of the ML algorithms in this specific crypto-mining problem [30]. Accordingly, it is expected that additional network attacks to the 5G Core using encrypted traffic will be detected by the Security Analytic Engine with new ML models based on the same aggregated data provided by the Security Data Collector.

5 Conclusions and Future Work

This paper presents a model-based approach to a data aggregation framework, relying on data modelling and network monitoring protocols, able to incorporate and pre-process data flows in multidomain environments, and suitable for closed-loop network management. The framework is based on standardized modelling mechanisms and architectures, and the initial version reported here has been integrated with open-source modules for metadata management tool and network orchestration. Using this initial version, the applicability of the framework to a realistic security issue in 5G networks has been proposed.

From this point, we plan to consolidate the framework, enhancing and better formalizing the metadata description mechanisms and making them suitable to accommodate the widest possible variety of data sources and consumers. Work is needed as well to evaluate the extensibility capacity of the framework and its supporting components, and to explore how it adapts to different network scenarios.

Acknowledgements

The research leading to these results received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 871808 (INSPIRE-5Gplus) and no. 856709 (5GROWTH). The paper reflects only the authors' views. The Commission is not responsible for any use that may be made of the information it contains.

References

[1] V.M. Hernández, R. Silva-Ortigoza: "Automatic Control with Experiments", Springer 2019. ISBN 978-3-319-75804-6

[2] ETSI GS NFV 002: "Network Functions Virtualisation (NFV); Architectural Framework", December 2104. Available online at https://www.etsi.org/deliver/etsi_gs/nfv/001_099/002/01.02.01_60/gs_nfv002v010201p.pdf

[3] ONF TR-502: "SDN architecture", June "014. Available online at https://opennetworking.org/wp-content/uploads/2013/02/TR_SDN_ARCH_1.0_06062014.pdf

[4] ETSI GS ZSM 009-1: "Zero-Touch Network and Service Management (ZSM); Closed-loop automation; Enablers", March 2021. Available

online at https://docbox.etsi.org/ISG/ZSM/Open/Drafts/009-1ed111_Cla_enab/ZSM-009-1_Cla_enabv0111.zip
[5] IETF RFC 6020: "YANG – A Data Modeling Language for the Network Configuration Protocol (NETCONF)", October 2010. Available online at: https://tools.ietf.org/html/rfc6020
[6] IETF RFC 6241: "Network Configuration Protocol (NETCONF)", June 2011. Available online at: https://tools.ietf.org/html/rfc6241
[7] W3C Recommendation: "Extensible Markup Language (XML) 1.0 (Fifth Edition)", November 2008. Available online at: https://www.w3.org/TR/2008/REC-xml-20081126/
[8] ECMA Standard ECMA-404: "The JSON Data Interchange Syntax", December 2017. Available online at: http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf
[9] Google Developers Resource: "Protocol Buffers Version 3 Language Specification", May 2020. Available online at: https://developers.google.com/protocol-buffers/docs/reference/proto3-spec
[10] IETF RFC 8040: "RESTCONF Protocol", January 2017. Available online at: https://tools.ietf.org/html/rfc8040
[11] Linux Foundation: "gRPC core concepts, architecture and lifecycle", 2020. Available online at: https://grpc.io/docs/what-is-grpc/core-concepts/
[12] ETSI GS CIM 006: "Context Information Management (CIM); Information Model (MOD0)", July 2019. Available online at: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/006/01.01.01_60/gs_CIM006v010101p.pdf
[13] ETSI GS CIM 009: "Context Information Management (CIM); NGSI-LD API", August 2020. Available online at: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.03.01_60/gs_CIM009v010301p.pdf
[14] ETSI SAREF: "SAREF: the Smart Applications REFerence ontology", 2020. Available online at: https://saref.etsi.org/core/v3.1.1/
[15] ETSI White Paper No. 31: "NGSI-LD API for Context Information Management", January 2019. Available online at: https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp31_NGSI_API.pdf
[16] IETF RFC 1157: "A Simple Network Management Protocol (SNMP)", May 1990. Available online at: https://tools.ietf.org/html/rfc1157
[17] Linux Foundation: "What is Prometheus?", 2020. Available online at: https://prometheus.io/docs/introduction/overview/

[18] IETF RFC 7011: "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", September 2013. Available online at: https://tools.ietf.org/html/rfc7011

[19] B. Chatras: "Applying a Service-Based Architecture Design Style to Network Functions Virtualization," 2018 IEEE Conference on Standards for Communications and Networking (CSCN), Paris, 2018. DOI: 10.1109/CSCN.2018.8581751

[20] ETSI GSM ZSM 002: "Zero-touch network and Service Management (ZSM); Reference Architecture", August 2019. Available online as: https://www.etsi.org/deliver/etsi_gs/ZSM/001_099/002/01.01.01_60/gs_ZSM002v010101p.pdf

[21] FIWARE Foundation: "Orion Context Broker Quick Start Guide", 2020. Available online at: https://fiware-orion.readthedocs.io/en/master/quick_start_guide/index.html

[22] ETSI OSM: "OSM Quickstart", 2020. Available online at: https://osm.etsi.org/docs/user-guide/01-quickstart.html

[23] Tabatabaei F., Wells D. (2016) "OSINT in the Context of Cyber-Security". In: Akhgar B., Bayerl P., Sampson F. (eds) Open Source Intelligence Investigation. Advanced Sciences and Technologies for Security Applications. Springer, Cham. https://doi.org/10.1007/978-3-319-47671-1_14

[24] D. Berman, A. Buczak, J. Chavis, C. Corbett: "A Survey of Deep Learning Methods for Cyber Security", 2019. Information, 10, 122, DOI: 10.3390/info10040122.

[25] J. Ortiz, R. Sanchez-Iborra, J. Bernabe, A. Skarmeta, C. Benzaid, T. Taleb, J.P. Wary: "INSPIRE-5Gplus: intelligent security and pervasive trust for 5G and beyond networks", August 2020. Proceedings of the 15th International Conference on Availability, Reliability and Security (pp. 1–10), DOI: 10.1145/3407023.3409219.

[26] S. Pastrana, G. Suárez-Tangil: "A first look at the Crypto-Mining Malware ecosystem: a decade of unrestricted wealth", 2019. Proceedings of the ACM Internet Measurement Conference IMC '19, pp. 73–86.

[27] J. Muñoz, J. Suárez-Varela, P. Barlet-Ros, "Detecting cryptocurrency miners with NetFlow/IPFIX network measurements", 2019. IEEE International Symposium on Measurements Networking, pp. 1–6, 10.1109/IWMN.2019.8804995.

[28] IETF RFC 3954: "Cisco Systems NetFlow Services Export Version 9", October 2004. Available online at: https://tools.ietf.org/html/rfc3954

[29] Telecommunication Networks Group - Politecnico di Torino: "Tstat, TCP STatistic and Analysis Tool", 2008. Available on line at: http://tstat.polito.it/

[30] A. Pastor, A. Mozo, S. Vakaruk, D. Canavese, D. Lopez. L. Regano, A. Lioy: "Detection of Encrypted Cryptomining Malware Connections with Machine and Deep Learning", 2020. IEEE Access, vol. 8, pp. 158036–158055, DOI: 10.1109/ACCESS.2020.3019658

Biographies

Antonio Pastor received the MSc. degree in industrial engineering from the Carlos III University of Madrid (UC3M), Spain, in 1999. Since then, he has been with Telefonica I+D (Research & Development), where he works on the engineering of different worldwide Telefónica networks. From 2006 to 2011 he has been working as an expert in network security for Telefonica Spain. Since 2012, as part of Telefonica Global CTIO, he leads innovation activities in network security based on network virtualization, SDN and artificial intelligence. He holds security certifications from ISACA and GIAC.

Diego R. López joined Telefonica I+D in 2011 as a Senior Technology Expert and is currently in charge of the Technology Exploration activities within the GCTIO Unit. Before joining Telefónica he spent some years in the academic sector, dedicated to research on network services, and was appointed member of the High-Level Expert Group on Scientific Data Infrastructures by the European Commission.

Diego is currently focused on applied research in network infrastructures, with a special emphasis on virtualization, data-driven management, new architectures, and security. Diego chairs the ETSI ISG on Permissioned Distributed Ledgers and the Network Operator Council of the ETSI ISG on Network Function Virtualization.

Jose Ordonez-Lucena received his B.Sc degree and M.Sc. degree in Telecommunications Engineering from the University of Granada in 2015 and 2017, respectively. He joined Telefónica I+D in 2018 as a Core & Platforms Technology Analyst, within the Global CTIO Unit. He is currently involved in technology exploration and innovative activities for 5G/B5G systems through different European research projects, with a focus on mobile

network architectures and end-to-end network slicing solutions, considering their applicability for public-private network integration scenarios. He also takes part in standardization activities, acting as Telefónica delegate in 3GPP SA5, ETSI ISG ZSM and GSMA 5GJA. From 2017, he is also pursuing a PhD in Telecommunications Engineering at the University of Granada.

Sonia Fernández received her B.Sc degree and M.Sc. degree in Telecommunications Engineering from the University of Cantabria in 2017 and from the Polytechnic University of Madrid in 2019, respectively. She currently works as Telecom Analyst in everis, an NTT Data Company. She joined Telefónica I+D in 2018 as a Transport & IP networks Analyst, within the Global CTIO Unit. She was involved in technology exploration and innovative activities for 5G systems through different European research projects, with a focus SDN networks and data-driven management, developing software tools to support network telemetry and autonomous management.

Jesús Folgueira joined to Network area in Telefonica I+D (Research & Development) in 1997. From 2017, he is the Head of Transport and IP Networks within Telefonica Global CTO unit, in charge of Network Planning and Technology.

He received his MSc degree in Telecommunications Engineering from Technical University of Madrid-UPM in 1994 and MSc in Telecommunication Economics in 2015 (UNED).

He is focused on Optical, Metro & IP Networks architecture and technology, network control plane (SDN) and Open Networking. His expertise includes Broadband Access, R&D Management, and network deployment.

He is an IEEE member and accredited professional in Telecommunications in his country.