Poster: IoT-Edge-Cloud Continuum data flow validation tool

Janis Judvaitis
Cyber-Physical Systems Laboratory
Institute Of Electronics and Computer Science (EDI),
Riga, Latvia
janis.judvaitis@edi.lv

Krisjanis Nesenbergs
Cyber-Physical Systems Laboratory
Institute Of Electronics and Computer Science (EDI),
Riga, Latvia
krisjanis.nesenbergs@edi.lv

Abstract

IoT-Edge-Cloud continuum is becoming a de facto for any IoT or AI-related service based on how the gathered data are processed and stored, with the new emergence of Edge computing. Since all of the citizens should take advantage of public services created in the IoT Edge-Cloud continuum the trust, security, and privacy questions are of utmost importance not only for the developers but also for other involved actors, such as owners, operators, and users. In this poster we propose a novel IoT-Edge-Cloud Continuum data flow validation tool aiming to assist the involved parties in aiming for the best possible outcome for society as a whole, contributing to the development of the IoT-Edge-Cloud continuum services as well as providing a mechanism for validation of data management mechanisms in theory as well as in practice.

1 Introduction

As the IoT-Edge-Cloud Continuum (IECC) becomes more popular the Internet of Things (IoT) and Artificial Intelligence (AI) are driving a new wave of digital transformation for our society. IoT devices such as cameras and sensors are increasingly used for public services, ranging from critical infrastructures (e.g., road monitoring and smart electricity grid) to healthcare services (e.g., remote patient monitoring for elderly people at home). IoT devices gather live data for advanced analysis and decision-making, usually powered by AI techniques. With the evolvement of IoT- and AI-powered public services, everyone potentially could be involved in various online activities without actively participating or even knowing it. Therefore, the trust and accountability of such services are critical for everyone in society, and the lack of trust in the current public services on security and privacy, including malicious intents of the service providers, is a major barrier to service innovation.

New paradigms and development tools emerge to enable a systematic and unified approach to IECC system development, for example, [4, 8], unfortunately, they are mostly focusing on improving the development part of the smart IECC systems leaving the crucial trust, security, and privacy management to the developers best knowledge and good intents. In this poster we present a IECC data flow validation tool aiming to fill the vertical integration gap between unified and streamlined smart IECC system development and enforced, verifiable, and accountable trust, security, and privacy policies by providing a clear view on the data flows of the whole IECC system. The aim of the proposed IECC data flow validation tool is to allow the system architects, designers, owners, and users to examine and validate how the data generated in the system are being processed, transformed, and transferred within the different computational spaces.

Similar approaches have already been proposed for other domains on a much smaller scale such as banking [7], web services [1], ontologies [2], etc. There are also tools available to system developers to help alleviate the complexity of IECC system design, such as NodeRed [5] or Digital Twin integration for IoT [3], bet they are more focused on the design and development part of the programming front, ignoring the data flow management.

2 IECC data flow validation tool

The proposed IECC data flow validation tool consists of four tightly coupled parts: (i) system-wide data flow policy, (ii) data flow management API, (iii) system-wide data access logging, and (iv) open source and verified plugins. Each of these parts is provided to the developers of the system in the form of software libraries ready to be used in their IECC projects.

System-wide data flow policy acts as the rule set placed upon the data flow within the IECC. The main purpose of the data flow policy is to mark different types of data and enforce their accessibility level within the system. In order for the proposed IECC data flow validation tool to be trusted for data flow verification there needs to be mechanisms put in place to prevent or at least report any tempering done to the data flow policy globally or locally.

Data flow management API is the sole enabler of the interaction with any data available in the IECC system. This is realized as an API available to the developers providing access to functions allowing to retrieve or provide data and

is enforced as the only input or output for each data processing instance in the IECC, keeping track of data origin and destination for each API call.

System-wide data access logging is self-explanatory, every request made for the data flow management API can be logged or saved for future evaluation according to data policy configured and then enforced. the gathered data can not only be used to verify that the system is behaving according to the rules by any authorized third party but also automatically observe any behavioral drifts or anomalies within the IECC system [6].

Open source and verified plugins are the main part of the system allowing for any interaction, investigation, or manipulation of the data to be injected in the IECC system to provide different functionality to different involved actors. Since plugins are also capable of interacting with the data, they should be verified and protected against any tampering, the verification ultimately should be entrusted to a competent third party, for example, the open-source community.

2.1 Plugins

The plugins operate in the "system" side of the IECC data flow validation tool therefore they theoretically have access to all the data available within the system. Data flow management API calls are used as entry points for any plugin interaction, restricting the scope of potential interaction. There are three distinct types of plugins we envision: (i) Validation, (ii) Manipulation, and (iii) Visualisation.

Validation plugins can validate the existing and defined data flows within the system, check if they adhere to any defined standards, for example, GDPR and similar, and flag potential security issues like plain text logging, etc. More advanced plugins could potentially evaluate the data privacy concerns by verifying the use of different data masking approaches, like Differential Privacy (DP) or leakage of private or sensitive data. It is also possible to use plugins intended to validate performance, but this is not the main focus of the proposed IECC data flow validation tool. The enforcement of system-wide data flow policies is also done through the plugin comparing the data flow transactions against the defined policies.

Manipulation plugins is a way to provide the IECC system developers with a secure and validated implementation of popular data management or masking algorithms, for example, DP, therefore not only accelerating IECC system development but also enabling a secure and trustable outcome.

Visualization plugins is designed to provide an easy-to-grasp visual representation of otherwise complicated internal data flows of IECC systems by concealing the non-relevant entities for the perspective depending on the necessary action. For example, a system developer, as well as a user or owner can use the envisioned example shown in Figure 1 to quickly get an understanding of the processing, transformation, and transmission of the gathered data and evaluate the flagged issues.

3 Preliminary Results & Next Steps

After the preliminary research, we have identified the necessity for IECC data flow validation tool and have formulated an initial approach for implementation, as well as secured two use cases related to critical infrastructure in order to validate the effectiveness and usefulness of the approach.

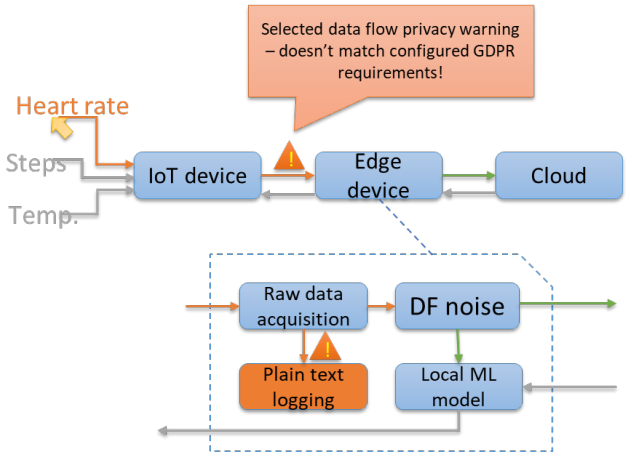

Figure 1. Example of visual data flow representation

The next steps are to implement the IECC data flow validation tool and validate it. For future research directions, we speculate that as the data flows within the system are available using data flow management API calls, it could be possible to add simulated data processing, providing or consuming instances working together with the real already implemented instances providing a unique opportunity not only to speed up implementation by using simulated parts of the IECC system but also evaluating different approaches by implementing them on more capable hardware before optimizing for IoT or Edge devices.

The authors are working towards the implementation of the proposed IECC data flow validation tool in order to validate and showcase the benefits of the approach.

Acknowledgements

This research is funded by the Latvian Council of Science, project "Smart Materials, Photonics, Technologies and Engineering Ecosystem" project No VPP-EM-FOTONIKA-2022/1-0001.

4 References

[1] C. Bartolini, A. Bertolino, E. Marchetti, and I. Parissis. Data flow-based validation of web services compositions: Perspectives and examples. *Architecting Dependable Systems V*, pages 298–325, 2008.

[2] J. Bārzdiņš, G. Bārzdiņš, K. Čerāns, R. Liepiņš, and A. Sprogis. Owlgred: a uml style graphical notation and editor for owl 2.

[3] G. Fortino and C. Savaglio. Integration of digital twins & internet of things. In *The Digital Twin*, pages 205–225. Springer, 2023.

[4] J. Judvaitis, R. Balass, and M. Greitans. Mobile iot-edge-cloud continuum based and devops enabled software framework. *Journal of Sensor and Actuator Networks*, 10(4):62, 2021.

[5] N. O'Leary and D. Conway-Jones. Node red-a visual tool for wiring the internet of things. *Retrieved July*, 4:2017, 2017.

[6] G. Rocher, S. Lavirotte, J.-Y. Tigli, G. Cotte, and F. Dechavanne. An iohmm-based framework to investigate drift in effectiveness of iot-based systems. *Sensors*, 21(2):527, 2021.

[7] S. Sadiq, M. Orlowska, W. Sadiq, and C. Foulger. Data flow and validation in workflow modelling. In *Proceedings of the 15th Australasian database conference-Volume 27*, pages 207–214, 2004.

[8] H. Song, R. Dautov, N. Ferry, A. Solberg, and F. Fleurey. Model-based fleet deployment in the iot–edge–cloud continuum. *Software and Systems Modeling*, 21(5):1931–1956, 2022.