FogProtect: Protecting Sensitive Data in the Computing Continuum*

Dhouha Ayed[1], Eva Jaho[2], Clemens Lachner[3], Zoltán Ádám Mann[4], Robert Seidl[5], and Mike Surridge[6]

[1] Thales, France
[2] Athens Technology Center, Greece
[3] TU Wien, Austria
[4] University of Duisburg-Essen, Germany
[5] Nokia Bell Labs, Germany
[6] University of Southampton, United Kingdom

Abstract. Computing resources are being moved towards the edge of the network, in the form of so-called fog nodes, providing benefits in terms of reduced latency, increased processing speed, data locality, and energy savings. Data produced in end devices like smartphones, sensors or IoT devices can be stored, processed and analysed across a continuum of computing resources, from end devices via fog nodes to cloud services. Data related to critical domains, such as healthcare, public surveillance or home automation, requires tailored data protection mechanisms, spanning the whole computing continuum.
The FogProtect project aims to provide novel advanced technologies and methodologies to ensure end-to-end protection of such sensitive data. Our generic solutions facilitate the provisioning and usage of applications and services in the computing continuum, by combining four technology innovations: (1) secure data container technology for data portability and mobility, (2) data-protection-aware adaptive service and resource management, (3) advanced data protection policy management, (4) dynamic data protection risk management models and tools.
The applicability and impact of those solutions is evaluated and demonstrated on three complementary real-world use cases in the area of (1) smart cities, (2) smart manufacturing, and (3) smart media.

Keywords: Fog Computing · Edge Computing · Data Protection · Privacy · Computing Continuum · Adaptive Systems · Policy Management.

1 Basic Project Information

The project "FogProtect: Protecting Sensitive Data in the Computing Continuum" is a research and innovation action funded by the European Union's Hori-

* This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871525 (FogProtect).

This paper was published in *Advances in Service-Oriented and Cloud Computing: International Workshops of ESOCC 2020*, pp. 179-184, 2021.

zon 2020 programme. The project runs from January 2020 to December 2022. The project website is at https://fogprotect.eu/.

The project consortium is led by Ubiwhere LDA (Portugal) and further comprises the following organizations: Athens Technology Center SA (Greece), IBM Israel Science and Technology LTD (Israel), University of Southampton (UK), Nokia Solutions and Networks GmbH & Co. KG (Germany), Thales SIX GTS France SAS (France), Technische Universität Wien (Austria), University of Duisburg-Essen (Germany), De Vlaamse Radio- en Televisieomroeporganisatie nv (Belgium).

2 Project Objectives

Cloud computing is transitioning from few large data centres to a truly decentralized paradigm, where resources are increasingly provided near the network edge, in the form of so-called fog nodes. Data produced in end devices (e.g., smartphones, sensors or other IoT devices) can be processed across a continuum of computing resources, comprising cloud services, fog nodes, and end devices [2]. By distributing data processing functionalities over this computing continuum, an optimal trade-off between conflicting goals – such as low network latency between data sources and data processors, high processing speed and low energy consumption – can be achieved [1].
However, this widely distributed processing of data also introduces new challenges concerning the protection of sensitive data [9, 10, 5]. During our research we identified the following major problems of data protection in the computing continuum [6]:

- Resource limitations of fog nodes and end devices constrain data protection methods
- Heterogeneity of fog nodes and end devices hampers consistent security
- Node connectivity meta-data can leak sensitive information, such as users' location
- Mobility of devices requires compliance with changing data protection policies
- Frequent changes at the edge imply highly dynamic changes of data protection risks
- Lack of transparency about stakeholders may lead to unauthorized data access

To tackle these issues, FogProtect will deliver new and advanced generic technologies, mechanisms, and solutions to ensure end-to-end data protection across the computing continuum. This also involves securing the whole life-cycle of data, taking into account the rights and obligations of data subjects, data controllers, data processors, and data users. In particular, FogProtect will make it easier for data controllers to comply with relevant data protection regulation, such as the EU General Data Protection Regulation (GDPR), and for data subjects to exercise the rights stipulated by the regulation.

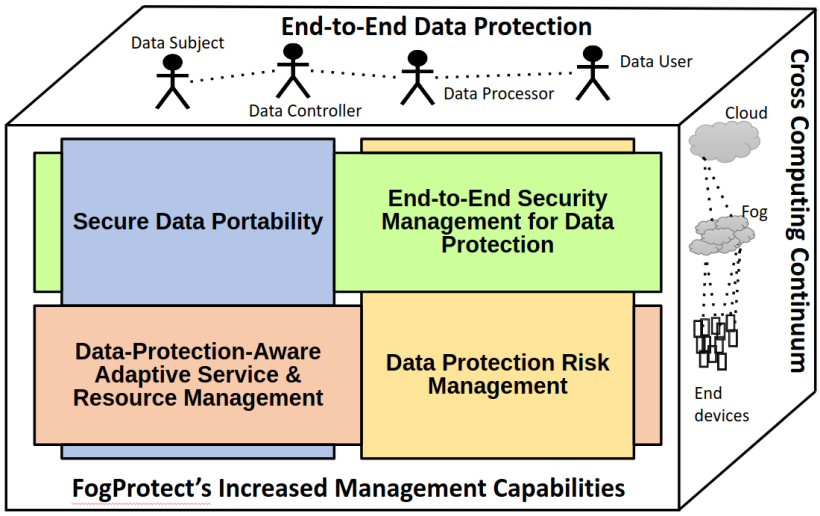

Fig. 1. An overview of FogProtects's four innovation areas.

As shown in Fig. 1, FogProtect combines four innovation areas to deliver increased management capabilities towards end-to-end data protection. A brief introduction of each innovation area is given in the next subsections.

2.1 Secure Data Portability

Dynamically composable, logical data encapsulation will provide data protection guarantees not only for the storage of data, but also for processing and egress to other sources in conformance with policies regulating the usage of data. To achieve that, FogProtect will create a layered framework, where different pluggable tools and technologies can be used at each layer in accordance with both the nature of the data and purpose of the data processing.

2.2 Data-Protection-Aware Adaptive Service and Resource Management

Based on continuous monitoring and analysis of data protection risks, appropriate adaptations will be carried out automatically. These adaptations will ensure the continued assurance of data protection in spite of changes [6]. For this purpose, FogProtect will create a model-based approach, in which the current system configuration in explicitly modeled in a machine-readable format, and kept up-to-date using monitoring.

For reasoning on possible adaptations to mitigate data protection risks, methods from artificial intelligence (optimization, planning, machine learning) will be used. Adaptations may relate to both the infrastructure of the computing continuum and applications running on that infrastructure. If multiple adaptations

are possible to mitigate the found data protection risks, FogProtect will aim to select the best adaptations concerning their impact on other goals like performance and costs.

Also, services in the computing continuum may initiate self-adaptations with the aim of improving performance or costs. FogProtect will ensure that the data protection implications of such adaptations are analyzed, and the adaptations are only performed if they do not lead to increased data protection risks.

2.3 Data Protection Policy Management

A data protection policy management solution requires being flexible and robust to support the orchestration layer of the fog architecture and support the big volume of interaction and data transfer between end-user devices, fog environment and cloud computing data centres. Therefore, we provide an end-to-end data protection policy management framework dealing with the distrusted, multi-tenant, and dynamic nature of fog nodes and instances.
The framework is based on a data protection policy definition formalism supporting the specification of security requirements related to fog nodes and instances and a system to orchestrate and chain data protection functions according to smart decision process based on dynamically interpreted security policies and data protection regulation.
The data protection policy formalism offers the administrators the flexibility to bound data of various subjects to a given application while having the ability to migrate across multiple fog nodes during the data lifecycle with a smart reasoning on policies and a refinement of the implication of data protection policies on various fog nodes and end-user devices.
The orchestrated data protection functions take into account the distributed nature of the environment and could be virtualized enforcement points, enforcement points for constrained environment, attack detection functions, etc.. For instance, the best practice for devices that deliver data on demand is to position a Policy Enforcement Point (PEP) closest to the data to protect. Consequently, various PEPs need to be deployed at several devices and nodes of the fog. However, such enforcement requires the knowledge of the current protection policy in force that is generally managed in a centralized point. The services running on constrained devices might not have a constant connectivity to a centralized decision point to enforce the policy. In this case, a lightweight enforcement process is needed. It can for example be based on a standalone access token without permanently relying on a central decision point.

2.4 Data Protection Risk Management

We identify, analyse and control sources of cyber-security risks (i.e., threats) over the lifecycle of applications in the computing continuum. This involves (i) the development of knowledge and inference methods to construct predictive models of potential risks prior to deployment and (ii) the use of information acquired (and mostly only available) after deployment, to diagnose run-time threats and

trigger adaptations or policy changes to manage the associated risks. The goal is to use an automated ISO 27005 risk assessment procedure [3, 4], embedding it into the autonomic management loop so risk levels can be taken into account. This will allow trading risk factors near the edge (such as limited physical protection) and near the data centre (such as the aggregation of data), as well as using data protection risk levels to constrain other autonomic management of factors such as cost, performance, energy use, etc.

3 Use Cases

Three complementary real-world use cases evaluate and demonstrate the applicability and impact of the FogProtect solutions introduced in the previous chapter. The use cases are focusing on different industrial sectors spanning multiple contexts: (1) smart cities, (2) smart manufacturing, and (3) smart media. Each use case entails different, specific data protection challenges from these sectors. This will enable being the ideal platform to reflect the whole life-cycle of data, taking into account the rights and obligations of data subjects, data controllers, data processors, and data users by ensuring end-to-end data protection across the computing continuum.

On the one hand, the use cases help the project team to identify requirements and constraints that the solutions developed in the project need to address. On the other hand, the demonstration and validation of the solutions developed in the project will also be carried out in the context of the use cases.

4 Current Project Status

The project has started recently. The focus of the first project phase has been on the elicitation of requirements from different sources (the project's use cases, the relevant literature, relevant standardization activities etc.) and on defining the architectural and technological foundations of the project, based on well established formalization methods, standards and processes [7]. Additionally, the state of the art is analyzed, regarding cutting-edge technology, concepts, and architectures in the area of application and infrastructure orchestration as well as data protection, risk management, and policy management.

The next step is the detailed specification of the technical components of FogProtect and their interfaces. A first prototype – integrated, tested, and validated on the use cases – will be available at month 18 of the project (June 2021). In the second half of the project, the FogProtect solutions will be refined and extended in a second iteration, taking into account the experience with the first prototype as well as new technical developments.

FogProtect leverages knowledge and experience gathered in previous work of project partners in the recently finished RestAssured project [8]. Therefore, we expect to be able to quickly resolve the architectural questions and to start working on the specific innovation areas.

References

1. Bellendorf, J., Mann, Z.Á.: Classification of optimization problems in fog computing. Future Generation Computer Systems **107**, 158–176 (2020)
2. Bonomi, F., Milito, R., Zhu, J., Addepalli, S.: Fog computing and its role in the internet of things. In: Proceedings of the first edition of the MCC workshop on Mobile cloud computing. pp. 13–16 (2012)
3. Chakravarthy, A., Wiegand, S., Chen, W., Nasser, B., Surridge, M.: Trustworthy systems design using semantic risk modelling. In: Proceedings of 1st International Conference on Cyber Security for Sustainable Society. pp. 49–81 (2015)
4. Goeke, L., Heisel, M., Mohammadi, N., Surridge, M.: Systematic risk assessment of cloud computing systems using a combined model-based approach. In: 22nd International Conference on Enterprise Information Systems, Prague, May 2020. p. to appear (2020)
5. He, T., Ciftcioglu, E.N., Wang, S., Chan, K.S.: Location privacy in mobile edge clouds: A chaff-based approach. IEEE Journal on Selected Areas in Communications **35**(11), 2625–2636 (2017)
6. Mann, Z.Á.: Data protection in fog computing through monitoring and adaptation. In: KuVS-Fachgespräch Fog Computing 2018. pp. 25–28 (2018)
7. Mann, Z.Á.: Notions of architecture in fog computing. Computing p. to appear (2020), https://doi.org/10.1007/s00607-020-00848-z
8. Mann, Z.Á., Salant, E., Surridge, M., Ayed, D., Boyle, J., Heisel, M., Metzger, A., Mundt, P.: Secure data processing in the cloud. In: European Conference on Service-Oriented and Cloud Computing. pp. 149–153. Springer (2018)
9. Roman, R., Lopez, J., Mambo, M.: Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges. Future Generation Computer Systems **78**, 680–698 (2018)
10. Stojmenovic, I., Wen, S., Huang, X., Luan, H.: An overview of fog computing and its security issues. Concurrency and Computation: Practice and Experience **28**(10), 2991–3005 (2016)