Data protection in fog computing through monitoring and adaptation*

Zoltán Ádám Mann
paluno – The Ruhr Institute for Software Technology
University of Duisburg-Essen
Essen, Germany

***Abstract*—Fog computing has enormous potential to offer increased computational capacity with acceptable latency near the network edge. However, fog computing also introduces many risks relating to the protection of sensitive data, which threaten the practical adoption of the fog computing paradigm. In this paper, the main challenges of data protection in fog computing and potential mitigation techniques are briefly reviewed. We argue that, given the highly dynamic nature of fog computing systems and the negative side-effects of existing data protection techniques, such techniques should be used adaptively, always in accordance with the relevant data protection risks. We sketch an approach to monitor a fog system and activate data protection techniques adaptively, followed by a research agenda to elaborate the details of the proposed approach.**

I. Introduction

Fog computing is the natural next step in the evolution of cloud computing, bringing cloud-like elastic compute capacity to the network edge, near to end user devices [1]. This way, computation-sensitive tasks can be offloaded from the end devices (like mobile phones, wearable devices, or cameras) to fog resources (i.e., compute resources at or near the network edge, e.g., in routers, base stations, or geographically distributed data centers of telecommunication providers). Offloading is advantageous for many applications that require higher computational capacity than what is available in end devices. Compared to offloading compute tasks to a large centralized cloud data center, fog computing has the advantage of considerably lower latency in the data transfers, which is essential for several time-critical applications [2].

Nevertheless, fog computing is also subject to several challenges. In particular, fog computing offers a plethora of opportunities for malicious parties to gain access to, or even manipulate, sensitive information [3]. Some of these threats are inherited from cloud computing, but some are new and specific to fog computing. More importantly, concerns about data protection can significantly hinder the adoption of the fog computing paradigm.

Of course, there are several known security techniques with which the access to sensitive data can be protected. However, the available techniques also have some limitations (e.g., some assume the availability of special hardware) or drawbacks (e.g., overhead). Therefore we argue that security techniques should be applied in an adaptive way. That is, the most appropriate technique should be selected based on the current situation. Adaptations should be carried out at run time, since also the situation may change dynamically at run time. Therefore, the current system state has to be monitored, so that risks concerning data protection can be identified and mitigated on the fly.

The contributions of this paper are as follows:

- A review of data protection challenges in fog computing;
- A proposed framework for adaptive handling of risks relating to data protection;
- Identification of research challenges to realize the proposed concept.

II. Data protection challenges in fog computing

For a more detailed survey of the general field of security and privacy in fog computing, the reader is referred to [4]. Here we only review the most important challenges related to the protection of sensitive information in fog computing.

Just like in cloud computing, users lose control of their data by uploading them to a server that is beyond their control [5]. Thus, the provider operating the given cloud or fog resource may get access to users' confidential data. The provider may also let third parties access the data – intentionally or unintentionally, with or without consent from the user – so that also these third parties may abuse the data. Moreover, because of the intrinsic multi-tenancy of both the cloud computing and fog computing paradigms, other users may also use the same server or fog resource, which might make it possible for those other users to gain unauthorized access to sensitive data. In some cases, it is also possible that users try to get access to confidential data of the provider, for instance to get to know important business secrets about the provider's infrastructure. All these types of attacks are conceivable in both cloud and fog computing.

In addition, there are some specific characteristics of fog computing that make data protection even more challenging than in cloud computing:

- **Reduced physical protection**. While cloud data centers are typically protected by strict physical access control mechanisms (e.g., doors that can be opened only by authorized personnel with their entry cards), fog resources are often deployed "in the wild" where malicious parties can get physical access much more easily. Even more

* This paper was published in: B. Koldehofe, A. Reinhardt, S. Schulte (Eds.): KuVS-Fachgespräch Fog Computing 2018, Technical Report, Technische Universität Wien, pp. 25-28, 2018

importantly, fog computing is mostly based on wireless networking technologies which may be broken into without physical contact. In contrast, cloud computing is mostly based on wired networks, which are easier to protect.

- **Less clarity about stakeholders**. In cloud computing, users choose service providers explicitly and deliberately, also giving explicit consent regarding the use of their data. On the other hand, in some fog computing scenarios, a device may use a variety of fog services for offloading computations, without the user of the device – or the data subject about whom the device is collecting data – being aware of the stakeholders that operate those resources or have otherwise access to the resources.
- **Direct access to confidential information**. A device using fog computing resources may leak sensitive personal information even without transferring any data explicitly to the fog resources. For example, since devices prefer to use nearby fog resources, an attacker might be able to determine a user's approximate location or the route of a mobile user based on which fog resources the user's device has connected to, thus violating location privacy [6], [7]. Another example is the violation of usage privacy in smart metering where the information gathered by smart meters reveals usage patterns of electronic devices in a household [8].
- **Scarce resources**. Existing methods for data protection, such as advanced cryptographic protocols or data obfuscation techniques, are often resource-intensive. This, however, is a problem in end devices that have limited computational capacity, limited battery power, and limited network bandwidth.

For these reasons, data protection is a very challenging problem in fog computing.

III. The case for adaptive data protection

Fog computing systems are very dynamic: end devices connect to fog resources and then disconnect, fog resources appear and disappear, wireless network connections get stronger or weaker etc. [9], [10]. With all those changes, also risk levels keep changing. For instance, from the point of view of an end device, risk levels may be low if the device can connect to a known and trusted fog resource, but the risk of data protection issues gets much higher if the connection to the trusted fog resource is lost and an unknown fog resource must be connected instead.

As already mentioned, existing security techniques that can ensure the protection of sensitive data often have downsides. For instance, homomorphic encryption makes it possible to offload computations on encrypted data to an untrusted server. Since the server gets access to the ciphertext only, it cannot abuse the actual data. However, homomorphic encryption introduces a large performance overhead [11]. Another option is the use of secure hardware enclaves, i.e., special hardware enabling the protection of code and data even from attackers with highest operating system privileges [12]. Performing computations in a secure enclave thus shields the data both from co-located applications and even from the operator of the server. However, also the use of secure enclaves incurs some overhead (although lower than in the case of homomorphic encryption) and even more importantly, it presumes the availability of appropriate hardware.

For these reasons, we argue that data protection mechanisms should be applied in an adaptive manner. In other words, data protection mechanisms should be activated only when needed to minimize their negative impact on resource consumption; moreover, from available alternative mechanisms always the most appropriate one should be chosen, taking into account the sensitivity of the data and the current configuration of the fog system.

In the following subsections, we review how adaptive application of data protection mechanisms can be achieved – first in general, and then focusing on the viewpoints of end users and fog service providers, respectively.

A. Enabling adaptations

The fundamental model underlying most adaptive systems is a control loop according to the MAPE (monitor, analyze, plan, execute) principle [13]. This principle can also be applied to the problem of adaptive data protection in fog computing, as follows.

The basis for any decision-making is monitoring. That is, the current configuration of the fog computing system needs to be monitored, including the available resources, the planned computations, the involved data, and any other information that may have an impact on risks (e.g., known vulnerabilities or reputation of stakeholders). Monitoring may be complemented by prediction, e.g. to predict the future availability of wireless network connections or the duration of offloaded tasks [14]. Based on the information provided by monitoring and potentially prediction, an analysis has to be carried out to determine the risks of data protection violation and the possible risk-mitigating actions. The results of the analysis form the input to planning. The aim of planning is to decide which risk-mitigating action(s) to take, based on the impact of the possible actions on both data protection risks and other system goals like performance or costs. Finally, the chosen actions have to be executed.

For implementing the MAPE loop, a model-based approach is advantageous. This means that a model of the fog computing system is maintained at run time in a machine-readable format. Monitoring updates the model so that it remains in sync with reality. Analysis and planning can be performed directly on the model, while execution ensures that modifications performed on the model are also transferred to the real world.

In our previous work, we have proposed a run-time model for reasoning about data protection in cloud systems [15]. This model should be extended and adjusted to capture the necessary entities of fog computing.

B. Adaptation in end devices

In the simplest case, an end device wants to offload some computations to a fog resource. Monitoring and analysis could

be used to answer the following questions:

- Can the provider of the fog resource be trusted (e.g., because of previous experience or because of high reputation in publicly available provider evaluations)?
- What security capabilities does the resource offer (e.g., secure hardware enclaves)?
- How sensitive are the data involved in the planned compute task?
- What would be the impact of the available client-side data protection mechanisms, and how critical would that impact be in terms of system goals like performance and energy consumption?

Based on these pieces of information, a sound decision can be made on the action to be taken. For example, if the data are not sensitive and/or the provider is trusted, then the computation can be offloaded without using further data protection mechanisms. Otherwise, if the targeted fog resource features secure hardware enclaves, then the computation should be performed within an enclave. Otherwise, if the resource situation allows it, homomorphic encryption should be used. If none of the above is applicable, then the computation should not be offloaded because the associated risks cannot be effectively mitigated.

There can also be more complicated cases, e.g., if not only an edge device and a fog resource are involved, but additionally a central cloud as well. To keep the analysis and planning manageable, a pattern-based approach can be used, like we suggested previously for cloud computing [16]. In this approach, the system configurations that would lead to unacceptably high risks of data protection violation are captured in the form of so-called risk patterns. If an instance of a risk pattern can be found as a substructure of the run-time model, then the risk is too high. An appropriate adaptation is needed so that the run-time model will not contain any of the identified risk patterns as substructure.

C. Adaptation in fog resources

For a provider of fog resources, the goal is to fulfill the data protection requirements, while serving as many end client devices as possible and also ensuring a smooth and efficient operation [17]. This leads to interesting optimization problems [18]. For example, if a subset of the resources owned by the provider offer secure hardware enclaves, then taking this into account when allocating client requests to resources is a useful lever for keeping costs low while fulfilling data protection requirements. Our previous experience has shown that, with appropriate optimization algorithms, if only a small fraction of resources offer secure hardware enclaves, this can already lead to considerable savings in energy consumption [19].

IV. Research challenges

In this paper, we sketched why adaptive data protection in fog computing is sensible and how it might be achieved. However, to make adaptive data protection in fog computing a reality, several research challenges need to be addressed (the list is not intended to be exhaustive):

- Appropriate models should be devised that incorporate all entities of fog computing deployments (along with their attributes and relations) that are relevant for data protection.
- The underlying trust models and attack models need to be better understood, categorized, formalized, and made available to automated run-time decision-making.
- A catalog of risk patterns needs to be elaborated that capture the relevant risks to data protection in fog computing.
- Appropriate monitoring techniques are necessary to efficiently and effectively monitor fog computing systems.
- Analysis, planning, and optimization algorithms need to be elaborated that can efficiently cope with decision-making on the different layers of fog computing systems (end devices, fog resources, cloud).
- Algorithm efficiency and scalability are crucial to ensure that adaptation works well in real time even under high dynamics.
- Testing, auditing, and verification techniques need to be devised to improve the credibility of data protection solutions in fog computing.
- The interplay among multiple autonomous systems that perform self-adaptations on their own needs to be better understood, including the possible coordination models and emergent behavior.

Moreover, it would be advantageous for fog computing research in general to define and make publicly available some representative examples that can be used to objectively assess and compare different approaches.

Acknowledgment

This work was partially supported by the European Union's Horizon 2020 research and innovation programme under grant 731678 (RestAssured).

References

[1] S. Yi, C. Li, and Q. Li, "A survey of fog computing: Concepts, applications and issues," in *Proceedings of the 2015 Workshop on Mobile Big Data*. ACM, 2015, pp. 37–42.

[2] O. Skarlat, S. Schulte, M. Borkowski, and P. Leitner, "Resource provisioning for IoT services in the fog," in *IEEE 9th International Conference on Service-Oriented Computing and Applications*. IEEE, 2016, pp. 32–39.

[3] I. Stojmenovic and S. Wen, "The fog computing paradigm: Scenarios and security issues," in *Federated Conference on Computer Science and Information Systems*. IEEE, 2014, pp. 1–8.

[4] S. Yi, Z. Qin, and Q. Li, "Security and privacy issues of fog computing: A survey," in *International Conference on Wireless Algorithms, Systems, and Applications*. Springer, 2015, pp. 685–695.

[5] C. A. Ardagna, R. Asal, E. Damiani, and Q. H. Vu, "From security to assurance in the cloud: A survey," *ACM Computing Surveys*, vol. 48, no. 1, 2015, article 2.

[6] Z. Riaz, F. Dürr, and K. Rothermel, "On the privacy of frequently visited user locations," in *17th IEEE International Conference on Mobile Data Management*, vol. 1. IEEE, 2016, pp. 282–291.

[7] T. He, E. N. Ciftcioglu, S. Wang, and K. S. Chan, "Location privacy in mobile edge clouds: A chaff-based approach," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2625–2636, 2017.

[8] A. Reinhardt, F. Englert, and D. Christin, "Averting the privacy risks of smart metering by local data preprocessing," *Pervasive and Mobile Computing*, vol. 16, pp. 171–183, 2015.

[9] A. Aral and I. Brandic, "Quality of service channelling for latency sensitive edge applications," in *IEEE International Conference on Edge Computing*. IEEE, 2017, pp. 166–173.

[10] S. Dräxler, H. Karl, and Z. A. Mann, "Joint optimization of scaling and placement of virtual network services," in *17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*. IEEE, 2017, pp. 365–370.

[11] W. Wang, Y. Hu, L. Chen, X. Huang, and B. Sunar, "Exploring the feasibility of fully homomorphic encryption," *IEEE Transactions on Computers*, vol. 64, no. 3, pp. 698–706, 2015.

[12] V. Costan, I. A. Lebedev, and S. Devadas, "Sanctum: Minimal hardware extensions for strong software isolation," in *USENIX Security Symposium*, 2016, pp. 857–874.

[13] J. O. Kephart and D. M. Chess, "The vision of autonomic computing," *Computer*, vol. 36, no. 1, pp. 41–50, 2003.

[14] G. Orsini, D. Bade, and W. Lamersdorf, "CloudAware: A context-adaptive middleware for mobile edge and cloud computing applications," in *IEEE International Workshops on Foundations and Applications of Self* Systems*. IEEE, 2016, pp. 216–221.

[15] Z. A. Mann, A. Metzger, and S. Schoenen, "Towards a run-time model for data protection in the cloud," in *Modellierung 2018*, I. Schaefer, D. Karagiannis, A. Vogelsang, D. Méndez, and C. Seidl, Eds. Gesellschaft für Informatik e.V., 2018, pp. 71–86.

[16] S. Schoenen, Z. A. Mann, and A. Metzger, "Using risk patterns to identify violations of data protection policies in cloud systems," in *13th International Workshop on Engineering Service-Oriented Applications and Cloud Services*, 2017.

[17] L. Wang, L. Jiao, J. Li, and M. Mühlhäuser, "Online resource allocation for arbitrary user mobility in distributed edge clouds," in *IEEE 37th International Conference on Distributed Computing Systems*, 2017, pp. 1281–1290.

[18] Z. A. Mann, *Optimization in computer engineering – Theory and applications*. Scientific Research Publishing, 2011.

[19] Z. A. Mann and A. Metzger, "Optimized cloud deployment of multi-tenant software considering data protection concerns," in *Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*. IEEE Press, 2017, pp. 609–618.