Privacy and Compliance in the Cloud-to-Edge Continuum

IBM Research Europe

Ambrish Rawat, Stefano Braghin, Beat Buesser,
Mark Purcell, Anil Kurmus, Andreas Wespi, Marc Stoecklin,
Kapil Singh, Arjun Natarajan, Juan Bernabé-Moreno

Motivation

It is undeniable that the rapid advancements of cloud computing have revolutionised IT services across many industrial sectors. However, with multiple service providers and the myriad of platforms available for use, the onus of infrastructure management is now with the application developers or the Dev/ML/SecOps engineers. While the developers can certainly benefit from AI-enabled automation or seamless interoperability via the cloud-continuum, there remain fundamental gaps in the management of privacy within these technologies. With the availability of compute at the edge and the rapid specialization of cloud capabilities, it's possible to bootstrap AI solutions that leverage the copious amount of data in distributed systems. However, the ease of application development can result in inadvertent omission of the required prudence for privacy and security concerns. With increasing decentralisation of cloud infrastructures, rapid specialisation of data sharing mechanisms and the rise of open-source, such challenges are even more pronounced. It is therefore crucial that advancements in technologies are compliant with the existing and upcoming regulations like GDPR [1], EU AI Act [2] and EU Digital Acts [3-5]. In this article we lay out some of the research challenges around privacy and accountability within multi-cloud technologies which we strongly believe need to be addressed for any responsible innovation and use of cloud technologies.

Current Status

Artificial Intelligence (AI) is a key driver for innovation in relation to a computing continuum and has been extensively investigated in projects pertaining to EU Horizon Calls ranging from those focussed on meta-OS to cognitive cloud. AI in this context often refers to the broad class of algorithms that help generalise predictions based on patterns recognised over large volumes of data. There are two dimensions to the interplay between AI technologies and cloud computing. First, is the development of AI applications that may be deployed or trained in cloud computing environments. Second, is the use of AI technology within the operational management of cloud computing. For instance, such algorithms are increasingly being developed for tasks like anomaly detection based on correlations obtained across metrics and logs. Other AI applications in cloud computing include fault localisation and root cause analysis that can operate in real-time and at scale [16]. These AI algorithms also benefit from the use of data sharing via collaborative approaches like federated learning [10, 11].

> As computation moves to the cloud, edge, or IoT devices, these become the new targets for attackers and makes it essential to adopt new defence strategies in a compliant fashion.

To enable the training and deployment of applications across a cloud-continuum it is essential to ensure interoperability across services when faced with heterogeneous data and compute infrastructure. Moreover, they need to adhere to the evolving regulatory requirements on data privacy. This leads to a classic struggle between utility and privacy risk for AI-based systems. One way forward is to ensure that the software services comply with relevant industry standards and regulations. But AI technologies are inherently statistical in nature and their stochastic dynamics make it challenging to establish such standards. This paves the way for research to advance technological development that strengthens end-user privacy within a cloud continuum and eases the management of security and regulatory compliance for application developers. Several reports show increasing costs for compliance (up to 30% [17]), in particular when dealing with data-heavy application, like AI-based solutions. Thus, any effort in streamlining compliance operation in cloud-to-edge continuum will significantly benefit the development of applications across industries.

Research Challenges

The promises of AI-based technologies for automation and optimization overshadow their vulnerability to privacy threats. The organisations building applications that handle sensitive data which can uniquely identify an individual, must take necessary measures to safeguard their systems against possible leaks. Furthermore, they need to achieve this within the specifications of different regulatory frameworks. Currently, this is achieved via siloed efforts from developers who use tools [14] that invoke the principles of privacy like minimization and anonymization for data processing and harness the guarantees from sophisticated approaches like advanced cryptographic techniques, and differential privacy to safeguard the AI approaches. This beckons the need for research that instead enables a cloud continuum to provide the necessary support and infrastructure to meet the management requirements of privacy,

information security and regulatory compliance.One must bear in the mind the perspective of different actors whose expertise could range from experienced end users to domain experts to system administrators or engineers or even auditors. The next generation cloud continuum needs to be armed with suitable abstractions for the diverse set of users to setup, build, deploy and use their solutions.

The operational management of a cloud continuum which leverages AI assistance and optimisations for performance improvements also faces the risks of privacy leakage. The regulatory aspects are perhaps even more pronounced for such systems as the different components of a cloud continuum may be hosted across different administrative and economic units or even geographies with different regulations, dictating the access and use of data. To consolidate these multilateral dimensions of privacy, we note the following broad themes as relevant for the next round of innovation within the cloud-edge-IoT space.

Equip the cloud continuum to help combat security and privacy risks of AI solutions

- AI solutions often require the use of sensitive data during training. Moreover, such data may be sourced from different regulatory domains and may be inherently heterogeneous.
- There exist numerous demonstrations of leaks in the form of inference attacks in AI or security vulnerabilities resulting from the multifaceted attack surfaces exposed via computing platforms.
- Currently the safeguarding is achieved manually by application developers with no direct support from the computing continuum.
- As data is moving across the cloud continuum, it gets more and more difficult for the end-user to control what data is flowing where and to avoid any possible compliance violation. Automated solutions are needed that manage data lineage and prevent data leaks.
- Recent advances in AI, program analysis and automated exploit generation techniques provide a unique opportunity to research novel and improved exploit prevention mechanisms for the large commodity open-source software commonly targeted by attackers, such as the Linux kernel.

Standardise the use of Privacy Enhancing Technologies (PETs) within a computing continuum to assist regulatory compliance

- Unclear adoption of PETs from a compliance point of view.
- Some guidelines have been presented [6, 7, 12, 13] but not globally accepted, leaving uncertainty for the application developers.
- AI-based and statistical-based approaches are currently underutilized because of a lack of generally accepted standards [8, 9]
- AI-based solutions are potentially shifting liability to service providers, but no clear legal interpretation has been reached.

Foster equitable and accountable development of capabilities in the continuum

- EU-based continuum should provide infrastructure to facilitate continuum consumption in accordance with current regulation by providing tools, services, and verifiable guarantee at a platform level.
- Promote the shift of compliance overhead from the users to the platform, thus enable faster adoption and exploitation by smaller organizations, and facilitate innovation in regulated environments.
- Platforms should be encouraged to participate in open markets, enabling fair and transparent exchange of data, services, resources and capabilities across users and providers.

References

[1] General Data Protection Regulation, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504
[2] Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (AI Act) and amending certain Union legislative acts: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206
[3] The Digital Services Act package, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package
[4] The Digital Services Act: ensuring a safe and accountable online environment, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en
[5] The Digital Markets Act: ensuring fair and open digital markets, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en
[6] https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2022/09/ico-publishes-guidance-on-privacy-enhancing-technologies/
[7] NIST Privacy Framework, https://www.nist.gov/privacy-framework
[8] IBM FactSheets Further Advances Trust in AI, https://www.ibm.com/blogs/research/2020/07/aifactsheets/
[9] NIST AI PROGRAM - AI Factsheet, https://www.nist.gov/system/files/documents/2023/03/30/AI%20Fact%20Sheet%200615%20FINAL.pdf
[10] From privacy to partnership: the role of Privacy Enhancing Technologies in data governance and collaborative analysis, https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/
[11] Federated learning key to securing AI, https://venturebeat.com/ai/federated-learning-key-to-securing-ai/
[12] National Strategy to Advance Privacy-Preserving Data Sharing and Analytics https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Strategy-to-Advance-Privacy-Preserving-Data-Sharing-and-Analytics.pdf
[13] Protecting Data: Can we Engineer Data Sharing?, https://www.enisa.europa.eu/news/protecting-data-can-we-engineer-data-sharing
[14] UN Handbook on Privacy-Preserving Computation Techniques https://unstats.un.org/bigdata/task-teams/privacy/UN%20Handbook%20for%20Privacy-Preserving%20Techniques.pdf
[15] UN Guide on Privacy-Enhancing Technologies for Official Statistics, https://unstats.un.org/bigdata/task-teams/privacy/guide/
[16] Bianchini, Ricardo, Marcus Fontoura, Eli Cortez, Anand Bonde, Alexandre Muzio, Ana-Maria Constantin, Thomas Moscibroda, Gabriel Magalhaes, Girish Bablani, and Mark Russinovich. “Toward ml-centric cloud platforms.” Communications of the ACM 63, no. 2 (2020): 50-59.
[17] Compliance Risk Study – 2022, Accenture https://www.accenture.com/nz-en/insights/consulting/compliance-risk-study