Privacy by Design Helps Blockchains Comply With GDPR

Abstract

Privacy concerns for personal data have significantly increased among internet users in recent years. Enormous amounts of data are being generated online that is used by third party organizations. The European Union has strictly mandated data regulation by formulating General Data Protection Regulation (GDPR) to ensure data regulation, and enhance transparency on the processing and storage of end-user data. In this paper, we evaluate and detail how permissioned blockchains, mainly used by enterprises, adhere to the GDPR's core principles. We bring out some non-trivial points on how permissioned blockchains inherently adhere to some of the GDPR principles concerning privacy, confidentiality and integrity of data and features such as 'Right to be forgotten'. The paper details some of the features of Hyperledger Fabric[1] such as private data collections, zero knowledge proofs, and private channels that comply with GDPR, and discusses the factors that influence the extent of compliance.

Introduction

Protection of personal data is a growing concern in the internet community. The Equifax Data Breach[2] and the Facebook Cambridge Analytica[3] debacle are two recent incidents that have exposed sensitive data of millions of US consumers. Personal data can reveal insights on customer, their thoughts, attitudes, and preferences. Reports suggest that identity theft is one of the most popular cybercrimes, because the more of our personal data that is available, the more easily can it be misused. Meanwhile, users lack awareness of the organizations' data usage and storage policies and hence, there is a pressing need to ensure stringent data privacy. The General Data Protection Regulation (GDPR)[4] has standardized data protection laws across the European Union, and has imposed rules on processing and controlling Personally Identifiable Information (PII). It is only with GDPR that regulatory compliance is sought in a more robust and mandated manner[5].

Privacy by Design[6] is the core principle of GDPR. Privacy by Design mandates that organizations must provide utmost priority to data privacy throughout the process of designing products/services that process personal data. Ever since GDPR came into force, organizations are now amending their applications to comply with its GDPR requirements.

With the emergence of Bitcoin, the underlying blockchain technology has been gaining popularity due to its decentralized nature and is being widely envisioned for application in various domains such as banking, fintech and supply chain. Initial versions of blockchain such as Bitcoin and Ethereum were public. However, as enterprise applications require greater accountability, permissioned blockchains became increasingly popular. Permissioned blockchains have verified identities and need members to run nodes, validate transactions, execute smart contracts, and read historic transactions. The level of decentralization needs to be mutually agreed upon by members of the blockchain network.

The core guarantees of GDPR aim at protecting user privacy by enforcing consents, data controls, data processing security, and other design guidelines. In this context, we evaluate the GDPR compliance of permissioned blockchains.

Consent on data collection and processing

Users are often prompted online with ambiguous agreements indicating data sharing with third parties that they agree to, without understanding the terms and conditions. GDPR attempts to strengthen the consent mechanism and enables lawful data processing with explicit consent by the subject.

This ensures only relevant and essential personal data shall be collected for legitimate purposes. The data subject has the right to know the purpose of processing, data disclosure details, and the duration of storage of collected data. It is mandatory that the processor processes the data only upon receiving controller instructions, assesses the impact, tracks processing activity, ensures security of processed data, and employs external processors only with controller consent.

While the requirement for user consent for data processing is the same in blockchain setting, the number of stake holders (or joint controllers) involved has increased. This has increased the level of complexity in ensuring compliance and accountability while maintaining user transparency.

Another principal GDPR mandate is the right to demand data erasure or withdraw consent subsequently. As data is replicated across multiple nodes in blockchain applications, data erasure is a complex process, one that is challenging for stakeholders to adhere to.

Privacy, confidentiality and integrity

Privacy and security in blockchain need to adhere to the following properties:

- Unlinkability of transactions
- Anonymity of users
- Confidentiality of transactions

Permissioned blockchains rely on cryptographic primitives like encryption and pseudonymous identities to adhere to these properties. While pseudonyms ensure unlinkability of submitted transactions and user anonymity, encryption guarantees authorized availability of transaction data. Permissioned blockchains facilitate access control mechanisms to further secure user data. Permissioned blockchains like Hyperledger Fabric (HLF) have private channels that can limit data within

nodes at the subnetwork level, to retain confidentiality. HLF also leverages novel cryptographic primitives called Zero-knowledge Proofs (ZKPs) which facilitate proof generation for statements without actually revealing the statement itself. The inherent immutable nature of blockchains ensures data integrity. Thus, permissioned blockchain models comply with GDPR mandates for data controllers.

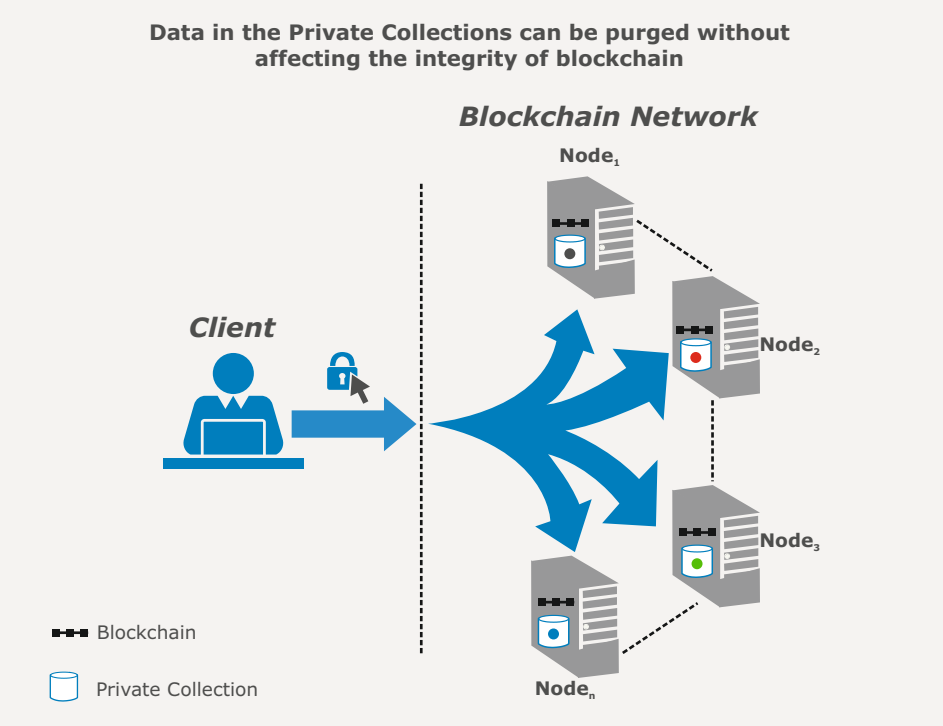

Figure 1: Purging data in private collections in blockchain

Data minimization and purpose limitation

Privacy by Design mandates only relevant (data minimization) and necessary (purpose limitation) data transaction, which is vital in ensuring data privacy across applications. Another crucial element of the GDPR is the 'Right to be forgotten' that regulates erasure obligations. This is a challenge, given that blockchain records transaction data permanently on the ledger.

To combat this limitation, permissioned blockchains like HLF have recently introduced a feature called Private Data Collections (in HLF v1.2.0)[7], which allows transaction data to be shared with a subset of members in the network (channel) as private data. This feature enables transaction data to be split into two – public and private. Private data can be put into different collections and shared with authorized peers within

the collection, as per pre-defined policy. Others receive a hash of the private data as evidence. This hash is stored in the ledger, along with the public transaction data. The private data may be purged manually without affecting the integrity of the blockchain. It can also be configured to be automatically purged after a certain duration (specified in terms of number of blocks – blocks to live).

However, this solution is GDPR compliant only assuming the parties involved have no malicious intent. Secure deletion is impossible if parties with malicious intent choose to copy the data. Formation of a consortium with clear, stringent rules is the only possible solution.

Organizations must ensure that the data is deleted from the production system, upon customer request for data deletion. However, data may need to be archived for historical, scientific research and audit purposes. Organizations need to be transparent with users about the need for data backup, and their plans for data security and deletion. Firms must evaluate between soft-delete and hard-delete options depending on the need to archive and incorporate application level access controls.

Chameleon Hash[8] may be used to enable 'Right to be forgotten'. This hash function comes with a trapdoor key to detect collision hash for the given input and data may be replaced with fictitious data. Chameleon Hash also enables 'Right to rectification' to rectify inaccurate data. However, a trapdoor key needs to be securely distributed among the nodes using a Multi-Party Computation (MPC) protocol.

Other issues

Data Protection Impact Assessment (DPIA) is a process to evaluate and minimize privacy risk in applications. As blockchain is still evolving, there is a need to formulate thorough assessment procedures for it to be effective. GDPR also defines that certifying bodies may attribute certifications to Data Controllers as GDPR compliant. However, setting standards for such certifications can be complicated for blockchain applications.

Future forward with compliant blockchains

Stringent regulations are being enforced to ensure data privacy, given the growing awareness on personal data protection among customers. GDPR is one such regulation that is mandated by the European Union to protect the personal data of European citizens. The extent to which individual blockchain platforms comply with GDPR differs across platforms, will be determined by platform features and the willingness of the involved parties to comply with the guidelines. According to Gartner, by 2023, over 25% of GDPR-driven, proof-of-consent implementations will involve blockchain technology. Even as regulatory bodies continuously update governing legislations to keep pace with evolving technology to protect user data and identity, businesses also need to do their bit to ensure data is safeguarded. There is mounting pressure globally to implement a holistic privacy management program and ensure that businesses proactively evaluate their data collection processes to maintain compliance and adherence to privacy regulations.

References

[1] https://www.hyperledger.org/projects/fabric

[2] https://www.equifaxsecurity2017.com/

[3] https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-scandal-everything-you-need-to-know.html

[4] https://eugdpr.org/

[5] https://www.computerworld.com/article/3273890/security/tech-talk-as-gdpr-looms-companies-rush-to-comply.html

[6] Cavoukian, Ann. "Privacy by design in law, policy and practice." A white paper for regulators, decision-makers and policy-makers (2011)

[7] https://hyperledger-fabric.readthedocs.io/en/release-1.2/private-data/private-data.html

[8] https://link.springer.com/chapter/10.1007/978-3-662-54388-7_6

About The Authors

Harika Narumanchi

Harika Narumanchi is a researcher in the cybersecurity and privacy research area at TCS Research and Innovation (R&I). Her research broadly focuses on applying cryptography and blockchain solutions to business-oriented scenarios. She graduated from the Jawaharlal Nehru Technological University, Hyderabad, India, with a master's degree in information technology, with a specialization in information security.

Nitesh Emmadi

Nitesh Emmadi is a researcher in the cybersecurity and privacy research area at TCS R&I. His areas of research include computations on encrypted data with a broader interest in application security, applied cryptography and blockchains. He looks closely into the practical side of novel systems and provides consulting services for evaluating and building products. Nitesh received his master's degree in information technology, with a specialization in information security, from IIIT, Hyderabad, India.

Contact

Visit the Research and Innovation page on www.tcs.com

Email: innovation.info@tcs.com

Blog: #Research and Innovation

About Tata Consultancy Services Ltd (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled, infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model™, recognized as the benchmark of excellence in software development. A part of the Tata Group, India's largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at www.tcs.com

TCS Design Services I M I 12 I 19