An Introduction to Blockchain's Growing Industrial Impact

Henry F. Korth, Fellow, IEEE

Abstract

At its core, a blockchain is a database, an organized collection of information. What makes blockchains distinctive is the special type of security they offer, both data security, and secure addition of new data by untrusting partners. The various blockchains in existence use various data structures and algorithms, but collectively they enable novel means for enterprises and individuals to access, share, and create data in ways that will disrupt industries and create new ones. No enterprise that uses information can escape permanently the impact of blockchain technology. In this short article, we shall explore some of these and note the mathematics, computer science, and systems engineering that underpins these trends. A longer exploration appears in [1] and [2, Chap. 26].

Blockchain data properties

The two data properties that characterize a blockchain database are *irrefutability* and *immutability*. Irrefutability arises from transactions being signed digitally by their submitter using public-key encryption [3, 4]. Immutability arises from two components of blockchain design. First, blocks link to their prior block along with a cryptographic hash of that prior block. The result is that if one were to modify one block all subsequent blocks would have to be modified. Therefore, holding only the hash of the most recent block allows one the ability to validate a blockchain. Secondly, protection from the most dramatic form of modification – deletion of data – is achieved by a high degree of replication. That replication itself creates a problem: the need to for all holders of a replica to reach consensus on each new block. Much of the criticism of Bitcoin is based on the energy-consumptive *proof-of-work* algorithm used to secure the chain. Consensus need not be energy intensive. Ethereum [5] is far along a transition plan to a higher performing and more efficient *proof-of-stake* algorithm. Other efficient algorithms include the *proof-of-history* algorithm in Solana [6], the decentralized trust applied in Stellar [7], and the mix of cryptography and randomization via *verifiable random functions* [8] in Algorand [9].

Blockchain in the enterprise

If a blockchain is *private*, that is, subject to admission control by some trusted authority, highly efficient consensus algorithms can be used since there is no need to guard against a malicious actor adding an arbitrarily large number of nodes to the network. However, in an enterprise setting, a traditional database is a viable, and often preferable framework for enterprise data. A traditional database typically outperforms a blockchain in terms of throughput as well as latency. *Trust* and *authority* are the key factors in identifying those parts of an enterprise database best managed on a blockchain. If all users of the system are managed by a single authority or if they all trust some single authority, then a traditional database is usually the right approach. However, if the enterprise database collects data from separately controlled organizations, as happens, for example, in a supply chain, then

the limited-trust model of a private blockchain has significant merit. In such a setting, there may be trust in a single organization for admission control into the private blockchain, but that organization may not be trusted in its handling of data should a controversy arise in the future. A simple example of this could be a supply chain for an end-user product that is later recalled for a defect. Various members of the supply chain may have a financial interest in trying to pass off responsibility to others. Suspicions or misunderstandings may arise. These are easily settled if the activity in the supply chain is documented by digitally signed transactions that were added to the blockchain by consensus of participants in the supply chain. Enterprise blockchains can be built on a public blockchain instead of a private one. Employing a public chain effectively outsources maintenance issues but incurs the possibly high costs of running transactions on a public chain.

Implementing a business relationship through a blockchain makes it possible to define many aspects of that relationship in code rather than in a legal contract. The unambiguous nature of code as compared to a natural language may reduce the number of post-agreement disputes. Enforcement actions can be implemented directly in code without waiting for some organization to file a claim or take other action. Of course, code may contain bugs. Careful attention to validation of code in a blockchain setting is essential. This may appear to be a drawback to a blockchain-based approach, but validation of legal contracts is equally essential and the many court cases over traditional contracts shows that use of English is at least as bug-prone as computer code.

Blockchain in finance

Finance is a hotbed of blockchain activity. Automated market makers and lending platforms are disintermediating traditional human-driven financial systems. Numerous central banks are contemplating the issuance of digital currencies backed by their national fiat currency. El Salvador has made Bitcoin legal tender. Although this topic falls in the domain of financial engineering, it is more removed from IEEE focus and so here we refer readers to a recent text on this topic [10].

Blockchain in infrastructure

All forms of infrastructure have become interconnected via the Internet. Blockchain extends the basic concept of interconnection to interoperability among untrusting systems. Systems that previously had a more command-and-control architecture are now becoming fully or partially decentralized, creating challenges in consensus and governance. While decentralized architectures appear to be a natural application for blockchain, the performance constraints (particularly as regards latency) raise a cautionary flag. On such example is the future electric grid, which will be increasing interconnected and which will include not only traditional power companies but also a wide array of “prosumers,” consumers who, at times, are also producers. An initial feasibility study in [11] addressed this issue for power-system sensor-data validation and concluded that a blockchain-based approach is feasible.

Blockchain in health

The growing interconnection of health providers, both through large healthcare-provider networks and the ubiquity of electronic health records, make the health domain a natural target for cross-enterprise blockchain-based data integration. The unique challenges here arise from the uniquely stringent privacy constraints surrounding health data and the challenge of identity management, particularly in the case of emergency care outside the patient's normal healthcare network. Health data must be encrypted for security, yet available in the case of unanticipated care. All forms of health care must be added to the patient's record with the providers of that care unambiguously identified. These requirements are simple to state but challenging to achieve. Here, we propose a summary of a framework that can meet these goals.

We begin with the principle that health data is the property of the patient, not the institution. Patient data are encrypted with the public key of the patient and thus decryptable only by the patient, who is the possessor of the corresponding private key. At a point of care, the patient makes the private key available (perhaps via a small hardware device so that data can be accessed even if the patient is incapacitated). Healthcare providers add to patient's data a justification for the data access itself, plus data on care provided, all signed digitally by the provider. Maintaining these data on a public, decentralized blockchain database ensures fast global data access with the security of public-key encryption (and signature). Because health data are voluminous (especially imaging data), the data themselves are stored off chain and secured on-chain cryptographically using a Merkle-tree data structure. *Merkle trees* [4] allow an off-chain data-storage system to provide data with a proof of validity. Only the root node of this tree (referred to as the Merkle root), need be on-chain. Off-chain data storage can be provided on conventional systems or in a fully decentralized manner as exemplified by Filecoin [12]. Finally, there is the matter of mapping identity in the real-world and identity in the blockchain world. Identity is the one area where some external authority is needed so that both patients and healthcare providers know that a specific blockchain ID corresponds to a particular person or institution. That mapping can be provided from a well-known (presumably government) online source. Such sources of external authority providing data to a blockchain application are referred to as *oracles*.

Conclusion

The examples explored here show that blockchain databases fill a critical need in multi-enterprise information systems. They allow reliable data sharing among organizations and individuals who trust each other at most partially. The decentralized, yet shared data framework that blockchains provide enables emerging applications, such as globally-accessible individualized health care, reliable multi-provider future energy grids, supply chains with fine-grain traceability of products and their components, and much more. Recent advances and continuing research in blockchain systems themselves are creating

blockchains with high performance and energy-efficiency, thus bringing the promise of blockchain concepts to industrial practice.

References

[1] O. Malekan, *The Story of the Blockchain: A Beginner's Guide to the Technology That Nobody Understands Paperback,* Triple Smoke Stack, 2018.

[2] A. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*, 7th edition. McGraw Hill Education, New York, NY, 2020.

[3] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Trans. Inf. Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[4] R. C. Merkle, "A digital signature based on a conventional encryption function," in *Adv. Crypt.* Springer Heidelberg, 1988, pp. 369–378.

[5] V. Buterin. "Ethereum: A next-generation smart contract and decentralized application platform". Original version, 2013. https://ethereum.org/en/whitepaper/. URL updated Feb 2022

[6] A. Yakovenko, "Solana: a New Architecture for a High Performance Blockchain," version 0.8.13, https://solana.com/solana-whitepaper.pdf, 2017.

[7] D. Mazières. "The Stellar consensus protocol: A federated model for internet-level consensus," Technical report, Stellar Development Foundation, https://www.stellar.org/papers/stellar-consensus-protocol 2016.

[8] S. Micali, M. O. Rabin, and S. P. Vadhan. "Verifiable random functions," In *Proc. 40th IEEE Symposium on Foundations of Computer Science*, 1979.

[9] Y. Gilad, R. Hemo, S. Micali, G. Vlachos, and N. Zeldovich, "Algorand: Scaling Byzantine Agreements for Cryptocurrencies," *ACM Symposium on Operating System Principles*, 2017.

[10] E. S. Prasad, *The Future of Money: How the Digital Revolution is Transforming Currencies and Finance*, Belknap Press, 2021

[11] A. G. Colaço, K. G. Nagananda, R. S. Blum and H. F. Korth, "Blockchain-based Sensor Data Validation for Security in the Future Electric Grid," *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference*, 2020

[12] J. Benet, et al., "Filecoin: A Decentralized Storage Network," Technical report, Protocol Labs, July 2017, https://filecoin.io/filecoin.pdf

BIOGRAPHY:

Henry F. Korth received his Ph.D. from Princeton University. He is currently a Professor in the Department of Computer Science and Engineering at Lehigh University, where he is a member of the Scalable Systems and Software Lab, and Director of the Blockchain Lab in the Center for Financial Services. He is a fellow of both the IEEE and ACM. His research interests are in database systems, particularly transaction processing and distributed systems, with a current focus on blockchain systems.