Towards Coordination-Free Replication and Access Control for Partition-Tolerant Decentralized Systems

Florian Jacob [1]

Abstract: In decentralized systems, access control decisions and enforcement are typically implemented via consensus, in order to mimic the centralized model where a single entity makes and enforces decisions. However, coordination-based approaches suffer from inherent non-availability under partition: Without being able to contact others, replicas cannot stay safe and live at the same time, they either need to stop providing service or proceed with potentially compromising safety. With a focus on replication and access control, we look at challenges and solutions from the growing class of theoretical concepts and practical Byzantine-tolerant systems that aim for autonomous replica decisions and thereby availability under partition by forgoing coordination.

At the other end of Byzantine fault-tolerant consensus in the coordination spectrum lies the land of Byzantine fault-tolerant, coordination-free autonomy. There, we find relevant practical systems for collaborative communication and data storage, like Matrix [Th23]. Its public federation has more than 100 million users and 100 thousand servers, and nation states operate private federations: e.g. in Germany, BWI operates the "BundesMessenger" for the public sector [BW24], and Gematik standardized the "TI-Messenger" in healthcare [ge24]. At the example of Matrix, we present our work on understanding and formalizing autonomous replication [Ja21] and access control under Byzantine faults [Ja20; JH23].

From the perspective of Byzantine fault tolerance in decentralized systems, coordination-free autonomous algorithms have very interesting properties: They remain available under partition, can tolerate an arbitrary number of Byzantine Sybill replicas, and avoid coordination as a scalability bottleneck [KH20]. The challenge is whether and how given problems can be cleverly expressed to be solvable by an autonomous algorithm, if solvable without coordination at all. The fundamental limit of autonomous algorithms is that they only can ensure confluent invariants: If every replica locally checks the invariant, it has to hold globally. While this rules out invariants like keeping account balances ≥ 0 or preventing doubled spending of a currency token, there is a strong community pursuing conflict-free replicated data types [Al23; Sh11] and coordination-free access control [RIP23; WBP16].

We present append-only event logbooks that capture immutable events and their partial-order causal relation as a common basis to derive complex replicated data types and access control decisions. As a design guideline to maximize fault tolerance and scalability for Byzantine decentralized systems, we argue for hybrid designs: system designers should strive to maximize autonomous system parts, and employ coordination and consensus only for the system's core that requires coordinated replica behavior to ensure safety and liveness.

[1] Karlsruhe Institute of Technology, Decentralized Systems and Network Services, Karlsruhe, Germany, florian.jacob@kit.edu, https://orcid.org/0000-0002-5739-8852

References

[Al23] Almeida, P. S.: Approaches to Conflict-free Replicated Data Types, 2023, arXiv: 2310.18220 [cs].

[BW24] BWI: IT Für Deutschland: BundesMessenger: Souveränität und Sicherheit und Freiheit. Freier Messenger für die öffentliche Hand. 2024, URL: https://messenger.bwi.de/bundesmessenger.

[ge24] gematik: TI-Messenger: Schnelle und sichere Echtzeit-Kommunikation auch im Gesundheitswesen, 2024, URL: https://www.gematik.de/anwendungen/ti-messenger.

[Ja20] Jacob, F.; Becker, L.; Grashöfer, J.; Hartenstein, H.: Matrix Decomposition: Analysis of an Access Control Approach on Transaction-based DAGs without Finality. In: Proceedings of the 25th ACM Symposium on Access Control Models and Technologies. SACMAT '20, Association for Computing Machinery, New York, NY, USA, pp. 81–92, 2020, ISBN: 978-1-4503-7568-9.

[Ja21] Jacob, F.; Beer, C.; Henze, N.; Hartenstein, H.: Analysis of the Matrix Event Graph Replicated Data Type. IEEE Access 9/, pp. 28317–28333, 2021, ISSN: 2169-3536.

[JH23] Jacob, F.; Hartenstein, H.: On Extend-Only Directed Posets and Derived Byzantine-Tolerant Replicated Data Types. In: Proceedings of the 10th Workshop on Principles and Practice of Consistency for Distributed Data. PaPoC '23, Association for Computing Machinery, New York, NY, USA, pp. 63–69, 2023, ISBN: 9798400700866.

[KH20] Kleppmann, M.; Howard, H.: Byzantine Eventual Consistency and the Fundamental Limits of Peer-to-Peer Databases./, 2020, arXiv: 2012.00472 [cs].

[RIP23] Rault, P.-A.; Ignat, C.-L.; Perrin, O.: Access Control Based on CRDTs for Collaborative Distributed Applications. In: The International Symposium on Intelligent and Trustworthy Computing, Communications, and Networking (ITCCN-2023), in Conjunction with the 22nd IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom-2023). 2023.

[Sh11] Shapiro, M.; Preguiça, N.; Baquero, C.; Zawirski, M.: Conflict-Free Replicated Data Types. In (Défago, X.; Petit, F.; Villain, V., eds.): Stabilization, Safety, and Security of Distributed Systems. Vol. 6976, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 386–400, 2011, ISBN: 978-3-642-24549-7 978-3-642-24550-3.

[Th23] The Matrix.org Foundation CIC: Matrix Specification v1.9, tech. rep., 2023.

[WBP16] Weber, M.; Bieniusa, A.; Poetzsch-Heffter, A.: Access Control for Weakly Consistent Replicated Information Systems. In (Barthe, G.; Markatos, E.; Samarati, P., eds.): Security and Trust Management. Lecture Notes in Computer Science, Springer International Publishing, Cham, pp. 82–97, 2016, ISBN: 978-3-319-46598-2.