Optimistic Execution in State-Machine Replication Systems

Alexander Heß [1] and Franz J. Hauck [2]

Abstract: State-machine replication (SMR) builds on consensus protocols to establish an ordering of client requests, which ensures that all replicas receive requests in a deterministic order. Traditionally, requests are executed after the request ordering is definitive. Throughout the years, however, different approaches have been proposed where requests are either tentatively executed while the consensus instance is still running, or speculatively before or even without undergoing a consensus protocol. The main goal of such optimistic approaches is to reduce the performance impact induced by the consensus protocol.

Keywords: State-Machine Replication, Speculative Execution, Tentative Execution

The concept of tentative request execution has already been outlined by Castro et al. with the publication of the PBFT consensus protocol [CL02]. The basic idea is that replicas can already execute requests before the final communication step of the protocol is completed, since the request ordering can be considered to be definitive with high probability. However, in case a view-change occurs before the consensus instance is finalized, the replicas have to discard the results and roll back to their latest checkpoint. A few years later another optimization has been proposed for the PBFT protocol, which leverages additional speculation on the client-side to further reduce the latency impact of the consensus protocol [We09]. *Zyzzyva* leverages speculation on both client- and server-side to further reduce the communication overhead between the replicas during normal operation [Ko09]. Here, clients are responsible for detecting inconsistencies between the replica's responses and to report these such that the system is able to recover.

While the main goal of the before mentioned approaches is to reduce the request processing latency, there are also approaches that leverage speculative execution to increase system throughput. Some of these approaches utilize application-specific knowledge to optimistically perform concurrent execution of non-conflicting requests [Ka10, Ka12]. If the speculative execution leads to diverging results, a fallback mechanism has to be invoked to restore a consistent system state. In contrast, there have been some proposal that leverage software transactional memory (STM) to perform optimistic concurrent execution, which do not have to rely on a separate fallback mechanisms [Zh16, Ge23].

This talk will provide an overview about different optimistic mechanisms that can be employed to improve performance of SMR systems under optimal system conditions. Additionally, it will discussed under which circumstances these mechanisms are actually beneficial, given that the recovery from failed speculation can be relatively expensive.

1 Ulm University, Institute of Distributed Systems, Germany,
alexander.hess@uni-ulm.de, https://orcid.org/0000-0001-6837-2861
2 Ulm University, Institute of Distributed Systems, Germany,
franz.hauck@uni-ulm.de, https://orcid.org/0000-0002-7480-9617

Bibliography

[CL02] Castro, Miguel; Liskov, Barbara: Practical byzantine fault tolerance and proactive recovery. ACM Trans. Comput. Syst., 20(4):398–461, 2002.

[Ge23] Gelashvili, Rati; Spiegelman, Alexander; Xiang, Zhuolun; Danezis, George; Li, Zekun; Malkhi, Dahlia; Xia, Yu; Zhou, Runtian: Block-STM: Scaling Blockchain Execution by Turning Ordering Curse to a Performance Blessing. In: Proc. of the 28th ACM SIGPLAN Annual Symp. on Princ. and Pract. of Par. Prog., PPoPP 2023, Montreal, QC, Canada, 25 February 2023 - 1 March 2023. ACM, pp. 232–244, 2023.

[Ka10] Kapitza, Rüdiger; Schunter, Matthias; Cachin, Christian; Stengel, Klaus; Distler, Tobias: Storyboard: Optimistic Deterministic Multithreading. In: 6th Workshop on Hot Topics in Syst. Dep. (HotDep). USENIX, 2010.

[Ka12] Kapritsos, Manos; Wang, Yang; Quéma, Vivien; Clement, Allen; Alvisi, Lorenzo; Dahlin, Mike: All about Eve: Execute-Verify Replication for Multi-Core Servers. In: 10th USENIX Symp. on Oper. Syst. Des. & Impl. (OSDI). USENIX, pp. 237–250, 2012.

[Ko09] Kotla, Ramakrishna; Alvisi, Lorenzo; Dahlin, Michael; Clement, Allen; Wong, Edmund L.: Zyzzyva: Speculative Byzantine fault tolerance. ACM Trans. Comp. Syst., 27(4):7:1–7:39, 2009.

[We09] Wester, Benjamin; Cowling, James; Nightingale, Edmund B.; Chen, Peter M.; Flinn, Jason; Liskov, Barbara: Tolerating Latency in Replicated State Machines through Client Speculation. In: 6th USENIX Symp. on Netw. Sys. Des. & Impl. (NSDI). pp. 245–260, 2009.

[Zh16] Zhao, Wenbing; Yang, William; Zhang, Honglei; Yang, Jack Y.; Luo, Xiong; Zhu, Yueqin; Yang, Mary; Luo, Chaomin: High-throughput state-machine replication using software transactional memory. J. Supercomput., 72(11):4379–4398, 2016.