High-Performance Consensus Mechanisms for Blockchains

Signe Rüsch
TU Braunschweig, Germany
ruesch@ibr.cs.tu-bs.de

ABSTRACT

The popularity of Bitcoin and other blockchain platforms has increased remarkably over the last few years. However, the mining process based on Proof-of-Work (PoW) needed to reach consensus leads to high energy consumption. This increases participation cost, leads to network centralization due to expensive, specialized hardware, and is problematic for the environment. An alternative to this PoW mining are Byzantine agreement protocols. Current Byzantine fault tolerant (BFT) protocols, however, are not suited for high numbers of participants, and therefore have to be adapted with regard to performance and security, e. g. by utilizing novel networking techniques and trusted computing.

KEYWORDS

blockchain, consensus algorithms, Byzantine fault tolerance, Intel SGX

1 INTRODUCTION

In recent years, the popularity of cryptocurrencies such as Bitcoin and Ethereum has increased remarkably. Bitcoin is regarded as the first successful realization of a digital currency that was widely adopted. The underlying concept is that of a blockchain: it consists of blocks linked together by including the hash of the previous block and containing ordered transactions. Through this, the blockchain is secure against manipulation. Two common use cases for blockchain technologies are cryptocurrencies, such as Bitcoin, and Supply Chain Management (SCM), the management of material, information, and services for product manufacturing. For cryptocurrencies, the number of users is not known in advance and there is no regulation of nodes entering or leaving the network. This setting is called a permissionless blockchain. For SCM, however, this regulation is necessary as it allows business partners to document transaction flows. Here, permissioned blockchains should be employed, i. e. blockchains where access is restricted and the participants are known and identifiable. How the next block in a blockchain is created and which transactions are included is determined through the consensus mechanism employed by the blockchain platform. The most prominent consensus mechanism, which is currently used by Bitcoin, is Proof-of-Work (PoW). Here, participants choose nonces and calculate hashes until the resulting hash is lower than a certain threshold. This PoW mining, however, is computationally intensive and also wasteful with regard to energy consumption, which is now estimated to be higher than that of the Republic of Ireland [9]. The increasing interest in Bitcoin mining led to the development of expensive, specialized hardware, which results in a centralization of mining power in the network. Since this is hurting the principles of the decentralized blockchain platforms, other alternative consensus mechanisms are considered. One is Proof-of-Stake (PoS) where the user's economic stake in the network instead of her computational power is the deciding factor on whether she can propose the next block. Another prominent alternative are Byzantine agreement schemes, in which a group of replicas tries to reach a consensus on the execution order and result of client requests although a subset of these replicas may behave arbitrarily faulty. As long as $3f + 1$ replicas behave correctly, the agreement scheme can tolerate up to f faulty nodes. The performance and security issues of BFT schemes in the blockchain setting will be the focus of this research proposal. We start by investigating BFT in permissioned settings: first for the Hyperledger Fabric blockchain platform [7], then in using novel network techniques such as RDMA to improve the performance of BFT protocols. Furthermore, we give an outlook on future work on improving the scalability and suitability of BFT protocols to later enable their deployment in permissionless blockchain settings.

EuroDW'18, April 23, 2018, Porto, Portugal
2018.

2 RELATED WORK

Permacoin [14] and Primecoin [17] are examples of cryptocurrencies that aim to better utilize the resources needed by PoW, by requiring users to offer storage space and by searching for large prime numbers, respectively. PoS will replace Ethereum's PoW mechanism with the Casper protocol [3]: it first introduces a hybrid PoW/PoS system, and in future versions, the PoW is supposed to be replaced by a more efficient mechanism.

Byzantine agreement schemes are considered especially well-suited for permissioned blockchains in which all participants are known. Well-known examples include Tendermint [11], Quorum [4], and Chain [10], which utilizes the Federated Consensus algorithm. BFT protocols, however, face several limitations when employed in permissionless blockchains: (i) they do not scale well with the number of participants and their performance degrades drastically for the targeted network sizes, and (ii) they are not yet widely established in this setting due to additional security issues, e. g. Sybil attacks [5] where an attacker participates using several identities. This way, an attacker might partake in the BFT protocol as multiple participants, thereby manipulating and corrupting the process in her advantage. Current approaches try to solve these issues in multiple ways. HoneyBadgerBFT [15] works with a fixed set of servers to run the consensus; however, this leads to centralization and allows an attacker to specifically target these servers. In Stellar [13], each participant forms a quorum of other users, thus creating a trust hierarchy, which requires complex trust decisions. Algorand [8] allows participants to privately check whether they are chosen for consensus participation, and requires only one message per user, thus limiting possible attacks.

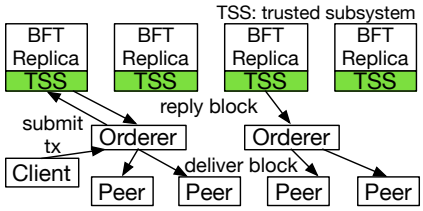

Figure 1: An HLF ordering service based on Hybster. Communication between BFT replicas is omitted for readability.

3 FIRST STEPS TOWARDS BETTER CONSENSUS

A prominent example for permissioned blockchain platforms is Hyperledger Fabric (HLF) [7]. During its re-design for the release of v1.0, a main design goal was extensibility. HLF v1.0 allows for multiple of its modules to be exchanged, such as the membership service or the consensus mechanism. As HLF is permissioned, this consensus mechanism is mainly responsible for receiving the transaction requests from the clients and establishing a total execution order. So far, these pluggable consensus modules include a centralized, single orderer for testing purposes and a crash-tolerant ordering service based on Apache Kafka [6]. This first orderer, however, offers a single point of failure, while the second offers only crash tolerance, not tolerance against Byzantine faults.

The Byzantine agreement scheme BFT-SMaRt was integrated into HLF v1.0 [2]. While this does offer protection against Byzantine faults, the results show that it also offers only limited scalability and throughput, i. e. the number of created blocks. The theoretical upper bound is determined by the cost of signature generation, in the presented measurements it is 86, 000 transactions per second. The current implementation also misses some crucial components, such as an access control mechanism or support for multiple blockchains shared between different users [2].

As our first step to investigate the behavior and performance of BFT protocols in permissioned blockchain platforms, we plan to integrate the Hybster protocol [1] into HLF v1.0, thereby offering protection against Byzantine faults at a higher performance. Hybster achieves this via exploiting the use of multicore CPUs and a trusted subsystem based on Intel SGX. The high-level architecture of our integration can be seen in Figure 1, which shows how a Fabric orderer receives a transaction, forwards it to the Hybster BFT cluster, and, after enough transactions to fill a block have been processed, creates and disseminates a new block to all connected peers via the orderers. To enable a transparent access to the BFT cluster for the orderers, we leverage a trusted proxy, called "Troxy" [12], to access the Hybster replicas. The Troxy also assumes responsibility for the voting on correct blocks. The created blocks can be cached in the Troxies connected to the replicas, all orderers can query any Troxy, and the peers can request created blocks from all orderers as well as exchange blocks via a gossip protocol. We therefore do not have to assume that orderers behave correctly to ensure that blocks are received by all peers and we also do not have to change the trust assumptions made by HLF.

The performance of BFT protocols can furthermore be improved by employing novel networking techniques. One possibility is Remote Direct Memory Access (RDMA), which enables machines inside a data center to access each other's memory in a zero-copy manner, without involving the operating system, thereby allowing for high throughput and low latency. As the BFT replicas in permissioned blockchains can be placed in data centers without compromising security, this is a feasible approach. We plan to change the communication layer of the Hybster prototype to use RDMA. Thus, the communication overhead between replicas should decrease due to eliminated data copying overhead on each machine, and allow for performance optimization by reducing the time necessary to reach consensus. While this integration requires adapting the application, it will not impact the protocol itself or lead to the abandonment of previously made performance-optimizing enhancements. Prior work has presented the employment of RDMA for state machine replication protocols [16].

4 SCALABILITY AND PERFORMANCE OF BFT PROTOCOLS

The scalability of BFT protocols with regard to the number of participants is highly limited and the performance of most protocols degrades with increasing number of involved replicas. For BFT-SMaRt in HLF, using more than 16 replicas leads to a performance decrease of up to $\approx 60\,\%$ for small messages. This effect is especially problematic for BFT deployment in permissionless blockchains.

The problem of BFT scalability is twofold: a high throughput (i. e. requests per seconds) as well as a large consensus group with good reconfigurability that can tolerate a high number of failures are both desirable properties in BFT protocols, but are often in direct conflict. Bitcoin mining, for example, supports thousands of participants, offers good reconfigurability, i. e. nodes can join or leave the network at any time, and can tolerate a high number of failures, while only processing a severely limited number of transactions per second. Most BFT protocols achieve a significantly higher throughput, but are limited to small groups of participants of ≤ 20 nodes and the group reconfiguration is not easily achievable [18].

Several approaches have been employed to remedy these problems, e. g. threshold cryptography, creating new consensus groups for every round, or limiting the number of necessary messages to reach consensus. Algorand [8], for example, scales up to 500, 000 users by employing Verifiable Random Functions, which are pseudo-random functions able to provide verifiable proofs that the output of said function is correct. In our future work, we aim to explore ways to improve the scalability of BFT protocols to enable their efficient deployment for permissionless blockchain environments.

Another important aspect is security: especially in permissionless blockchains, an adversary might launch several attacks, e. g. DoS or Sybil attacks. We plan to investigate how trusted computing, e. g. based on Intel SGX, can be leveraged to prevent this, for example by ensuring the uniqueness of a CPU in a network and therefore increasing the complexity for an adversary to create multiple identities without considerable cost.

5 SUMMARY

Nowadays, current BFT protocols deployed in blockchains still face severe limitations with regard to performance and scalability. Some improvements to scalability, e. g. those employed in HoneyBadgerBFT, introduce new attack scenarios or centralization in the network, or require the user to make complex decisions as in Stellar [13]. In our work, we will focus on several avenues of improving BFT performance and scalability, and their applicability as well as security for deployment in blockchains.

REFERENCES

[1] Johannes Behl et al. "Hybrids on Steroids: SGX-Based High Performance BFT". In: *Proceedings of the Twelfth European Conference on Computer Systems*. EuroSys '17. 2017.

[2] Alysson Bessani et al. "A Byzantine Fault-tolerant Ordering Service for the Hyperledger Fabric Blockchain Platform". In: *Proceedings of the 1st Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers*. SERIAL '17. 2017.

[3] Vitalik Buterin et al. *Casper the Friendly Finality Gadget*. 2017. URL: http://arxiv.org/abs/1710.09437.

[4] JPMorgan Chase & Co. *Quorum Whitepaper v0.1*. 2016. URL: https://github.com/jpmorganchase/quorum-docs/blob/master/Quorum%20Whitepaper%20v0.1.pdf.

[5] John Douceur. "The Sybil Attack". In: *Proceedings of 1st International Workshop on Peer-to-Peer Systems (IPTPS)*. 2002.

[6] Apache Software Foundation. *Apache Kafka*. URL: https://kafka.apache.org/.

[7] The Linux Foundation. *Hyperledger Fabric*. 2017. URL: https://hyperledger.org/projects/fabric.

[8] Yossi Gilad et al. "Algorand: Scaling Byzantine Agreements for Cryptocurrencies". In: *Proceedings of the 26th Symposium on Operating Systems Principles*. SOSP '17. 2017.

[9] Alex Hern. *Bitcoin's energy usage is huge – we can't afford to ignore it*. 2018. URL: https://www.theguardian.com/technology/2018/jan/17/bitcoin-electricity-usage-huge-climate-cryptocurrency.

[10] Chain Inc. *Chain Protocol Whitepaper*. 2017. URL: https://chain.com/docs/1.2/protocol/papers/whitepaper.

[11] Jae Kwon. *Tendermint: Consensus without Mining*. 2014. URL: https://tendermint.com/static/docs/tendermint.pdf.

[12] Bijun Li et al. "Troxy: Transparent Access to Byzantine Fault-Tolerant Systems". In: *Proceedings of the 48th International Conference on Dependable Systems and Networks*. DSN'18. Accepted for Publication. 2018.

[13] David Mazières. *The Stellar Consensus Protocol: A Federated Model for Internet-level Consensus*. 2015.

[14] Andrew Miller et al. "Permacoin: Repurposing Bitcoin Work for Data Preservation". In: *Proceedings of the IEEE Symposium on Security and Privacy*. S&P '14. 2014.

[15] Andrew Miller et al. "The Honey Badger of BFT Protocols". In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. CCS '16. 2016.

[16] Marius Poke et al. "DARE: High-Performance State Machine Replication on RDMA Networks". In: *Proceedings of the 24th International Symposium on High-Performance Parallel and Distributed Computing*. HPDC '15. 2015.

[17] *Primecoin*. 2014. URL: http://primecoin.io/.

[18] Marko Vukolić. "The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication". In: *Proceedings of the IFIP WG 11.4 Workshop – iNetSec 2015*. 2015.