Blockchains
Subtitle

Guilherme João Bidarra Breia Lopes

Master's disseration planning
Engenharia Informática
(2nd degree cycle)

Supervisor: Prof. Doctor Simão Melo de Sousa

janeiro de 2023

Palavras-chave

Inserir palavras-chave

Contents

Acronymms

NFT	Non-fungible token

Chapter 1

Introduction

1.1 Intoduction

Blockchains have gained significant attention in recent years for their potential to revolutionize various industries by providing a secure and transparent method for storing and transferring information and assets. At their core, blockchains are decentralized and permissionless ledger systems that leverage cryptography, consensus algorithms and State Machine Replication to maintain a shared history of all transactions on the network.

The increasing popularity of blockchains can be attributed to the growing demand for secure and transparent systems in various industries, particularly finance, supply chain management, digital identity and recently digital assets with the use of Non-fungible tokens.

In addition, the distributed nature of blockchains enables them to operate in a trustless environment, reducing the risk of single point of failure and ensuring the integrity of data stored on the network.

Consensus algorithms play a crucial role in the functioning of blockchains, as they determine the process by which transactions are validated, new blocks are added to the chain, and other details such as the selection of the next block producer.

Different consensus algorithms offer varying levels of security, scalability, and decentralization, and choosing the right algorithm for a specific use case is critical for the success of a blockchain network.

The main objectives of this dissertation are to compare different consensus algorithms for blockchains, evaluate the trade-offs between security, scalability, decentralization, and explore the potential for consensus algorithms to be plugged and changed on demand. Additionally, the dissertation will cover the implementation and testing of a blockchain network with a pluggable consensus algorithm and a thorough examination of the Tezos(REFERENCIA) blockchain as a testbed for consensus algorithms.

This dissertation will provide valuable insights into the field of blockchains and consensus algorithms, contributing to the development of more secure, scalable, and efficient blockchain systems, such as tools and methods to test consensus algorithms, to develop algoritms and to include them in already developed blockchain nodes.

1.2 Motivation

Despite the growing popularity of blockchains, there is currently a lack of tools and methods for easily swapping and testing different consensus algorithms in a live environment. This makes it difficult to determine which algorithm is best suited for a specific use case, particularly in the context of the blockchain trilemma, where it is challenging to achieve optimal

balance between scalability, security, and decentralization.

Additionally, developing consensus algorithms can be a complex and time-consuming task, and there is currently no standardized method for describing and implementing them. This presents a significant barrier to the adoption and evolution of blockchains, as it limits the ability of developers to experiment with new and innovative consensus algorithms. Also, the decentralization being one of the strenghts points of blockchain networks falls in the hands of a few selected developers.

This dissertation aims to address these challenges by providing a comprehensive analysis of different consensus algorithms for blockchains, and exploring the potential for consensus algorithms to be plugged and changed on demand. The results of this study will serve as valuable feedback for the development of a domain-specific language for describing consensus algorithms, and will provide insights into the most effective methods for testing consensus algorithms in a live environment.

The importance of this research cannot be overstated, as the ability to easily swap and test consensus algorithms is critical for the continued growth and evolution of blockchains. This will not only benefit the academic community but also industry stakeholders, who will be able to make informed decisions about which consensus algorithm is best suited for their specific use case.

In conclusion, this dissertation is motivated by the growing need for a comprehensive analysis of consensus algorithms for blockchains, and the desire to explore new and innovative methods for swapping and testing consensus algorithms in a live environment. The objective of this study is to contribute to the development of a standardized method for describing and testing consensus algorithms, ultimately advancing the field of blockchains and improving the security, scalability, and decentralization of blockchain networks.

1.3 Problem Statement

(REFERENCIA) (TODO: REMOVER REPETIÇÃO DAS COISAS) Consensus algorithms, being a fundamental aspect of blockchain technology, are subject to a range of emerging problems. Hardforks, network Denials of Service attacks, centralization of networks, and other relevant issues are challenging the stability and security of these systems. The choice between scalability, security, and decentralization, as described by the Blockchain Trilema, can lead to trade-offs in the design of consensus algorithms.

One of the challenges of consensus algorithms is the difficulty of changing the rules of the game after a blockchain network has been established. The hard forks of the Bitcoin, Bitcoin Cash, and Bitcoin SV networks have highlighted this issue. Similarly, the Ethereum network has also undergone several hard forks, leading to a split in its community.

Scalability issues are another challenge faced by some blockchain networks, including Bitcoin and Ethereum, which have struggled to accommodate increasing demand. In addition, some networks, such as Solana, have experienced outages, further highlighting the need for a more robust and reliable consensus system.

This dissertation aims to address these challenges by exploring the development of con-

sensus algorithms and ways to test them. The Tezos blockchain provides a suitable platform for this work, as it has a pluggable consensus system and allows for the implementation of different consensus algorithms. The goal is to learn how to develop consensus protocols in Tezos, develop the main entry points for a possible connection with a Domain Specific Language, and develop ways to test consensus algorithms. The first subject of such tests will be a proof-of-work consensus algorithm.

The question we want to answer in this dissertation is the following: How can we make a blockchain network that is flexible in terms of consensus algorithm selection and provides a platform for testing and comparison of different consensus algorithms? By addressing this question, we aim to provide a solution that will allow blockchain networks to be more scalable, secure, and decentralized, thereby addressing the challenges posed by the trilema. Additionally, this dissertation will serve as input/feedback for later work on a Domain Specific Language that will be used to describe consensus algorithms, making it easier to develop and test consensus algorithms in the future.

1.4 Document Organization

Chapter 2

Core Concepts and State of the Art

2.1 Core Concepts

(TODO: Mudar ”For better understanding....”) For a better understanding, this chapter describes the fundamental concepts of the topic of this thesis, which revolves around Blockchains, Consensus Algorithms, and State Machine Replication. A consensus algorithm refers to the method used by a network to reach agreement on a set of values in a distributed system. Blockchains, on the other hand, are append-only ledgers that are linked to each other by cryptographic hash functions, making them immutable. Blockchains can be considered as State Machine Replication as they store the current state of a system, and the consensus algorithm ensures that every copy of the ledger is updated consistently. The FLP theorem, Partial Synchrony, and BFT algorithms are related to the topic of consensus algorithms, and will be explained in this section. This thesis uses the Tezos blockchain as a case study, due to its ability to easily swap and upgrade its consensus algorithm, making it a suitable platform to study these topics.

State Machine Replication

(TODO: Adicionar aos acrónimos) (TODO: Verificar o que eu disse) (TODO: Falar do CAP System e fazer referencias) State Machine Replication (SMR) is a fundamental concept in decentralized systems that plays a crucial role in ensuring the consistency and availability of shared data. SMR works by representing the state of a system as a finite-state machine, which provides a clear and concise representation of the current state of the system. This representation allows for the use of consensus algorithms, which coordinate updates to the state and ensure that all nodes in the system agree on the current state. This leads to a consistent view of the state across all nodes, ensuring the data remains consistent even in the presence of node failures or network partitions.

It provides several key advantages in distributed systems. One of the main advantages, like mentioned befire, is that it can provide both consistency and availability, making it a popular choice for use in blockchain technology, databases, file systems, and other decentralized systems. Additionally, SMR provides mechanisms for recovery from node failures, ensuring that the state of the system can be recovered and can continue to operate. This is important in systems where the loss of a single node can have serious consequences for the entire network.

SMR also has the ability to handle network partitions and ensure that all nodes can continue to operate and reach consensus, even if the network is temporarily disconnected. This is accomplished through the use of consensus algorithms, such as Paxos, Raft, and Byzantine Fault Tolerance (BFT) algorithms, this last one being explained in a later section (TODO:

adicionar referencia à outra subsecção). These algorithms provide the mechanism for ensuring that all nodes agree on the current state of the system, even in the presence of network partitions or node failures.

In terms of performance, SMR can be designed to provide high performance and scalability, making it a suitable choice for large-scale distributed systems. This is achieved through the use of efficient algorithms and the optimization of network communication and data storage. Additionally, SMR provides mechanisms for ensuring the security of the state of the system, such as digital signatures and encryption, to prevent tampering or unauthorized access.

Overall, the concept of State Machine Replication is a fundamental component of decentralized systems that plays a crucial role in ensuring the consistency and availability of shared data (in case of Blockchains, to ensure that every node have the same chain). Its ability to provide both consistency and availability, along with its performance and security benefits, make it a popular choice for use in a variety of distributed systems, including blockchain technology, databases, file systems, and more.

Blockchain Data Structure and Networks

Blockchain is a data structure that is often compared to a linked list, since every node of the list points to a node in the list. In the case of, every block in the chain points to a previous block in the chain.

It is used to record transactions and is based on the concept of state machine replication. Each block in a blockchain contains a collection of verified transactions and a reference to the previous block in the chain. This creates a chain of blocks, each one linked to the previous one through a unique hash generated using a cryptographic function. That is, every block, except for the first block (sometimes called “Genesis” block), contains the hash of it’s previous (or parent) block. Once a block is added to the blockchain, it cannot be altered or deleted, making the ledger tamper-resistant and immutable, since, in order to change a block’s information, one would need to change every child’s block hash. One can also say that, the older the block, the more “immutable” it is, or the bigger the number of child blocks a block has, the harder is the tampering of such block.

Blockchain networks operate in a decentralized manner, with no central authority or entity controlling the network. The network reaches consensus on the contents of a new block through the use of consensus mechanisms, such as Proof of Work or Proof of Stake(TODO: Adicionar referencia) (These mechanisms will be explained in detail in later subsections).

The blockchain is a distributed ledger, with all nodes in the network having a copy of the ledger, ensuring availability and transparency.

State machine replication ensures that all nodes in the network have the same chain, which is critical in a decentralized system. The consensus mechanism in a blockchain network ensures that all nodes agree on the state of the system and that all nodes have the same chain. This helps to ensure the consistency and integrity of the data stored in the blockchain.

The blockchain can be considered the state of a machine, where copies of the state of this machine is stored in multiple nodes, and newer states, that is, the appending of a new block to the chain, is replicated to every machine.

Beyond cryptocurrency, blockchains have a wide range of potential applications in various industries, such as supply chain management, voting systems, and identity verification. The transparent nature of blockchains ensures that all transactions are publicly accessible and verifiable. The cryptographic functions used in blockchains, such as hashing and consensus mechanisms, provide a high level of security to the ledger and its transactions.

In conclusion, blockchain is a data structure that leverages the concepts of state machine replication to provide a secure and tamper-resistant ledger for recording transactions. The decentralized nature of blockchains and the use of consensus mechanisms ensure that all nodes in the network have the same chain, providing both availability and consistency.

Consensus

Consensus algorithms play a crucial role in decentralized systems, serving as the engine that drives State Machine Replication. These algorithms are used to reach agreement on the state of a system among all the nodes in a network, ensuring that all participants have a common understanding of the system's current state. The use of consensus algorithms has become widespread in various decentralized systems, such as blockchains, distributed databases, and distributed file systems, where it is crucial to maintain the consistency and integrity of data.

There are several important theorems related to consensus algorithms, including the FLP Impossibility Theorem and the CAP Theorem. The FLP Impossibility Theorem states that, in an asynchronous network, it is impossible to achieve both reliability and consensus in the presence of process failures. This theorem highlights the trade-off between reliability and consensus in decentralized systems. On the other hand, the CAP Theorem states that it is impossible for a distributed system to simultaneously provide consistency, availability, and partition tolerance. These theorems provide limits and trade-offs in decentralized systems, and serve as a useful reference for designers and researchers in the field.

(TODO: Por imagem) One subject of much discussion is the topic of the "Blockchain" trilemma, where people try to define the consensus of a blockchain based on a location in a area formed by a triangle.

The trilemma of blockchain networks refers to the trade-off between scalability, security, and decentralization. In other words, it is impossible to achieve all three goals at the same time in a blockchain network, and these trade-offs are the points of the trianle mentioned before. Scalability is often refered as Speed or Velocity.

For example, Bitcoin is a highly decentralized and secure blockchain, as there are many nodes (Around 40 thousand (TODO: Add citation)), yet lacks on scalability, since, the creation of a block takes 10min and the a transaction takes about 1 to 1.5 hours to complete (TODO: Add citation).

This trilemma is related to the CAP theorem, which states that in a distributed system, it is impossible to simultaneously guarantee consistency, availability, and partition tolerance. Consistency means that all nodes see the same data at the same time, availability means that all nodes can access the data at any time, and partition tolerance means that the system continues to work even if communication between nodes is lost.

In the case of Bitcoin, for example, increasing the scalability of the network would require

reducing the number of nodes that validate transactions, which would compromise the security and decentralization of the network. Similarly, improving the security of the network by adding more nodes would increase the complexity and reduce the scalability of the network.

Therefore, when designing a blockchain network, it is important to understand the trade-offs between scalability, security, and decentralization and to make trade-offs based on the specific goals and requirements of the network.

Consensus protocols, such as RAFT, Paxos, and BFT (Byzantine Fault Tolerance), are used to reach consensus in decentralized systems. These protocols have a set of formal requirements, including agreement, weak validity, strong validity, and termination. Agreement states that all correct processes must agree on the same value, while weak validity states that the output of each correct process must be the input of some correct process. Strong validity requires that all correct processes output the same value if they receive the same input value, and termination requires that all processes eventually decide on an output value.

Blockchains rely on consensus algorithms to maintain the integrity and consistency of their data. One of the most well-known consensus algorithms in blockchains is Nakamoto consensus, which is used in the original Bitcoin blockchain and relies on proof of work. Another popular consensus algorithm is Ethash, used by Ethereum, which is a form of proof of work. Proof of Stake is another consensus algorithm that uses the stake of a node to determine its ability to validate transactions and add new blocks to the chain. Delegated Proof of Stake, used by Tezos, allows nodes to delegate their validation power to other nodes, increasing the efficiency of the network while maintaining decentralization. Blockchains can also use hybrid consensus algorithms, combining elements of proof of work and proof of stake.

It is important to study various types of consensus attacks, such as 51% attacks, double-spending, and more, and how consensus algorithms are designed to prevent these attacks. Other consensus algorithms worth mentioning include Delegated Byzantine Fault Tolerance, Practical Byzantine Fault Tolerance, and Directed Acyclic Graphs. These algorithms have unique features and trade-offs, and the choice of a consensus algorithm for a particular blockchain will depend on its specific requirements and goals.

In conclusion, consensus algorithms are a fundamental component of decentralized systems, serving as the engine that drives State Machine Replication. These algorithms are used to reach agreement on the state of the system among all nodes in a network, ensuring that all participants have a common understanding of the current state of the system. The choice of a consensus algorithm will depend on the specific requirements and goals of a decentralized system, and the trade-offs between reliability and consensus, as well as availability and consistency, must be considered.

Bibliography