Sharding in Blockchain: A Scalability Solution

Dharaneeswar Reddy Rami Reddy[1], Madhana Mohit Konduru[2]

Department of Computer Science [1,2]

Email: DharaneeswarReddy.RamiReddy01@student.csulb.edu

Madhanamohit.konduru01@student.csulb.edu

Abstract:

Blockchain has started gaining popularity in many industries for providing secure, transparent, and trustless transactions. Blockchain is well known for its immutability and decentralized nature. However, we need to address the pressing issue of blockchain's scalability. As the size of the blockchain network increases, its ability to process and handle transactions quickly becomes a huge obstacle. One of the solutions to tackle this problem is Sharding. This paper aims to explore sharding in blockchain, various techniques in Sharding, and how it helps the scalability issue. We also look at the concepts of blockchain, its limitations, and how Sharding comes into play. By exploring the real-world applications of sharding in blockchain, we aim to establish the power of sharding in blockchain. By giving an in-depth overview of how sharding works, its challenges, and its contribution to blockchain, the paper aims to highlight the critical importance of sharding in the blockchain.

Keywords: Blockchain, Decentralized, Scalability, Sharding.

1. Introduction :

Definition of blockchain: Blockchain is a shared, immutable ledger that facilitates the process of recording transactions and tracking assets in a business network. An *asset* can be tangible (a house, car, cash, land) or intangible (intellectual property, patents, copyrights, branding). Virtually anything of value can be tracked and traded on a blockchain network, reducing risk and cutting costs for all involved.[1]

The data in a blockchain is stored in the form of data blocks which are connected cryptographically. That is each block holds a unique hash value which is connected to its adjacent block's hash value. Thus forming an immutable chain as a minute change in one block hash affects the entire chain. Algorithms like Proof of work (PoW) and Proof of stake(Pos) are utilized to add new blocks to the blockchain. All the nodes in the blockchain have a copy of the blockchain and the transactions made ensuring the transparency and security of the blockchain. The picture below describes the architecture of a traditional blockchain.

1.1 Challenges in Blockchain:

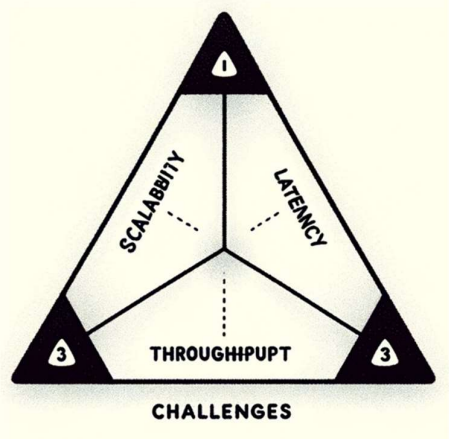

Fig. 1*: Challenges in Traditional Blockchain*

1.1.1 Scalability:

As the size of blockchain increases the ability of the blockchain to maintain and process transactions deteriorates. This is because in a blockchain each node has to maintain the record of the entire blockchain which leads to an issue in scalability.[16]

1.1.2 Throughput:

In blockchain for transactions to process every node needs to give its consensus and label it as a valid transaction. As the volume of transactions increases it gets difficult and time-consuming to process each transaction hence decreasing the throughput of the blockchain.

1.1.3 Latency:

As the blockchain network size increases the network congestion causes delays in getting the confirmations from every node thereby increasing the latency in the network. This is a major issue as latency can cause serious drawbacks when blockchain is used in a real-time application.

1.2 Why is scalability difficult to achieve:

There are six main challenges we have to overcome if want to achieve scalability. Firstly transaction bottlenecks- as more users join the blockchain network the amount of transactions can overwhelm the system, pose significant delays, and put a strain on the system. Blockchain is essentially an energy intensive system especially those that rely on proof of work. Thirdly with each transaction, there is pressure building up on the storage system which poses a threat of centralization. Every node in the Blockchain system has to process transactions as well as maintain a record this can cause significant network congestion. Blockchain mainly relies on the consensus mechanism which makes it difficult to find an agreement when there are huge numbers of nodes often making it slow. Lastly, blockchain poses many interoperability issues when interacting with other blockchains which needs to be addressed.

Why we need to address scalability challenges:
As technology advances and more and more people start using blockchain, eventually the blockchain will stagnate. To overcome this, we need to address issues like scalability which will increase the blockchain size depending on the traffic. Increasing the throughput also helps the blockchain handle the growing number of transactions. Lastly, if we want seamless and uninterrupted services reducing the latency should be our priority.

2. Sharding: A Scalability Solution:

Sharding has gained popularity as one of the best solutions to our pressing blockchain problem – Scalability. In blockchain, sharding refers to a technique that divides the network into smaller partitions, called shards, to improve scalability and increase transaction speed.

Sharding is a database management concept that has been around long before its application in blockchain technology. Its can be traced back to traditional database system where it was used to enhance database management and performance. In blockchain sharding emerged as a response to the scalability challenges faced by early blockchain networks like Bitcoin and Ethereum. Its introduction made a significant shift from the traditional method of every node validating every transaction. The adoption of sharding in blockchain networks like Ethereum 2.0 represents a huge step, showcasing the technology's ability to adapt and overcome inherent limitations.

In more technical terms, sharding is a form of database partitioning, also known as horizontal partitioning, where the database is horizontally partitioned into pieces (shards) to provide high concurrency and short response times. In the case of blockchain sharding, the network is split into multiple shards in which transaction validation occurs (instead of on the entire blockchain network), and each shard holds a unique set of smart contracts.[2]

2.1 Types of sharding:

Sharding in blockchain means dividing nodes into multiple shards and then dividing the activities and objects, such as transactions, blocks, and ledgers, into independent parts. The division of nodes enables them to process and store transactions that are only related to those specific shards.

There are three types of sharding in blockchain:

Network Sharding: In this type of sharding the entire blockchain network is divided into smaller shards or committees. Instead of verifying all the transactions in the network the nodes only receive and process a portion of the entire transaction.[3][15]

State Sharding: This means that instead of all nodes storing the entire historical ledger, each committee stores only the ledger associated with its shard.[3]

Transaction Sharding: In this type of sharding the transactions are divided into multiple sets and are assigned to different committees. Each committee reaches a consensus within its shard, reducing the computational load on individual nodes.[3]

These sharding types are interrelated but might not be adopted simultaneously. For example, in the early version of Ethereum's sharding required nodes to join two types of networks: one for propagating consensus-related information within the shard and another for transmitting transactions and blocks across the entire network, thus performing transaction and state sharding.[3]

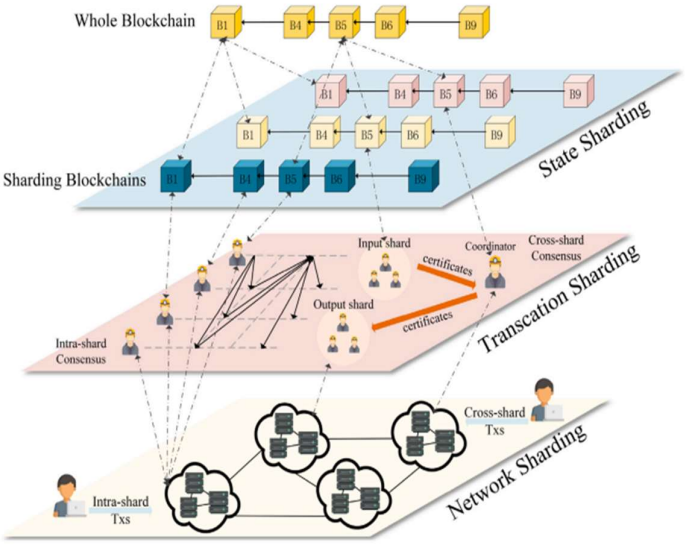

Fig. 2: *Relationship between types of Sahrding.[3]*

2.2 Decomposition of Sharding in Blockchain:

Let's take a look at how shards are decomposed into functional components and how each component works.

Node Selection: This component helps in selecting the nodes that will participate in the sharding process. It is important for maintaining the network's integrity and security.[1][4]

Epoch Randomness: This is the randomness used in the process of sharding, particularly for epochs which are specific periods during which sharding configuration remains unchanged.[4]

Node Assignment: After nodes are selected they are assigned to specific shards. This is an important step in the distribution of workload and responsibilities among different parts of the blockchain.[4]

Intra-Shard Consensus: This component deals with obtaining consensus within a shard. Each shard operates somewhat independently and reaching a consensus within each shard is essential for the overall functioning of the blockchain.[4]

Cross-Shard Transaction Processing: This involves managing transactions that span across different shards. It is a complex process since it requires coordination between shards to ensure the integrity and consistency of the blockchain.[4]

Shard Reconfiguration: Over time shards may need to be reconfigured for various reasons such as to balance load or respond to changes in the network. This component handles such reconfigurations.[4]

Motivation Mechanism: This component is related to the incentives and rewards for participating in the network including how nodes are motivated to maintain the integrity and performance of the blockchain. [4]

2.3 Sharding communication :

The Fig. 3 illustrates the architecture of a sharded blockchain device. It indicates three separate shards (Shard 1, Shard 2, Shard 3) functioning as a mini-blockchain in the larger blockchain network. Let's go through the key principles depicted in the picture:

Shard: Each 'Shard' represents a portion of the nodes within the blockchain community. Each shard is answerable for processing a fragment of all transactions, thereby distributing the burden throughout the network. [4]

Shard Member: These are the nodes which can be a part of a particular shard. They work collectively to validate and process transactions within their shard.[4]

Block: Each coloured block represents a set of transactions that have been established and introduced to the blockchain within a shard. The colour differentiation indicates the nature of the facts dealt with by every shard.[4]

Intra-Shard Communication: This is the conversation that takes place within a shard. Shard participants communicate with every different to attain a consensus on transactions and to validate them. The consensus algorithm best suited to the shard is used here.[4]

Consensus Algorithm: This is the process utilized by nodes inside a shard to agree on the state of the blockchain. It ensures that each node has an identical replica of the ledger.[4]

Coordinator: This entity coordinates the activities within a shard, which may additionally include the initiation of consensus rounds or the distribution of responsibilities among shard contributors.[4]

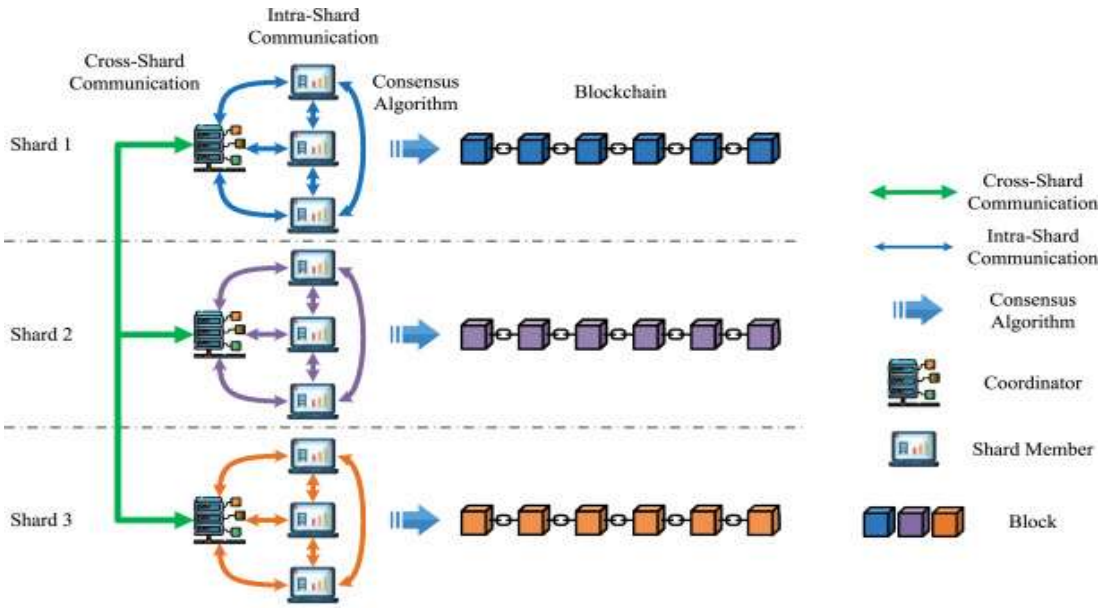

Fig. 3: *Sahrding Communication[4]*

Cross-Shard Communication: This represents the process by means of which shards communicate with each other. This is necessary when transactions have an effect on a couple of shard or whilst shards need to synchronize their states to make certain the integrity of the overall blockchain.[4]

Blockchain: This represents the complete chain of blocks from all shards. The blockchain is the whole set of all transactions processed by means of the community, assembled in a linear collection.[4]

In this example, every shard operates semi-independently but coordinates with other shards to hold a coherent and synchronized blockchain. Transactions are processed parallelly in different shards, and through cross-shard communication, the consequences are integrated into the overarching blockchain. This architecture permits for extended throughput and scalability, as each shard can procedure transactions concurrently with others, therefore improving the network's usual capability.[4]

2.4 Benefits of Sharding:

Although sharding provides multiple advantages lets look at the main ones:
At the top is **'Parallel Power**.' Sharding empowers the blockchain with parallel processing capabilities, similar to multi-threading in computing. Each shard operates concurrently, processing transactions and smart contracts simultaneously. This not only multiplies the transaction throughput but also accelerates the overall network performance.
The second advantage is **'Economical Operation**.**'** Sharding optimizes the use of resources across the network. By reducing the load on individual nodes, it cuts down the operational costs and lowers the barriers to entry.[17]
By spitting the blockchain into smaller shards, each node is not responsible for processing the entire network's **transactional load**. This causes the computational and storage workload to be spread out enhancing the scalability of the system.[17]
The fourth advantage is the **Efficiency in Large scale networks**. Sharding is particularly effective in large-scale networks. For example Ethereum is exploring sharding as part of its scalability solutions, aiming to significantly scale the network and enable it to process more transactions per second.
Finally, we have 'Democratized Participation.' Sharding democratizes the blockchain network by allowing more participants to join without the need for prohibitive computational power or storage capacity. It levels the playing field, providing opportunities for a wider community to contribute to and benefit from the network.

2.5 Challenges and concerns of sharding:

Let's take a look at the challenges and concerns of sharding. Sharding isn't just an incremental update; it's similar to **'Redefining Architecture**.' introducing sharding into existing blockchains demands significant architectural changes while ensuring the network remains operational.

The task of **'Synchronizing Fragments'** presents another complex challenge. When transactions span multiple shards, they must be orchestrated to maintain a consistent ledger. This level of coordination is comparable to managing multiple teams working on a collaborative project, each with its deliverables that must align perfectly with the others.

Then there's the imperative of **'Guarding the Fort.'** The introduction of sharding opens up new vulnerabilities in the network's security landscape. Each shard, while a fortress in itself, may introduce potential weak points that could be exploited. Maintaining the integrity of a sharded blockchain requires vigilance and the development of cutting-edge security protocols to protect against these new threats.
Efforts to tackle these challenges:

Advanced Consensus Mechanisms: "An efficient sharding consensus algorithm for consortium chains," by Xiaoxiong Wu et al. This paper proposes a new clustering-based sharding consensus algorithm (KBFT) that aims to address the issues of system communication congestion and reduced throughput in large-scale consortium chains. [5]

Security Enhancements for Shards: "On the Security and Performance of Blockchain Sharding," by Runchao Han, Jiangshan Yu, Haoyu Lin, Shiping Chen, and Paulo Esteves-Veríssimo. This paper conducts a comprehensive evaluation of blockchain sharding protocols, identifying new attacks, questionable design trade-offs, and open challenges[6].

3. Real-world Applications:

Sharding is a revolutionizing approach to tackle network enhancement, scalability and efficiency. Below are some real-world implementations of sharding in blockchain:

Ethereum 2.0: One of the popular examples of Blockchain Sharding is Ethereum's Beacon Chain which is a part of the Ethereum 2.0 update. It assigns validators to committees and chooses proposers for each slot using a procedure known as RANDAO. The Beacon Chain oversees the operation of Ethereum's proof-of-stake protocol across all of its shard chains including crosslink processing, validator storage and cross-shard transaction facilitation[7][8].

Polkadot: Polkadot uses a technology called Parachain for distributed database sharding. Parachains are simpler blockchain forms that attach to the security provided by a relay chain. This setup allows for secure message passing between parachains and independent computations within each parachain. This structure also allows the creation of parachains for efficient data storage and transaction operations which can further have their own parachains creating a tree like structure for distributed computations.[8]

Aspect	On-Chain Transactions	Off-Chain Transactions	Sharding in Blockchain
Definition	Transactions processed and recorded on the blockchain, requiring consensus for validation.	Transactions occurring outside the blockchain, enabling faster and cheaper transfers.	A method of dividing a blockchain into smaller segments (shards) to improve performance.
Advantages	Secure, transparent, immutable, support for smart contracts.	Scalable, lower costs, better privacy, flexible.	Increases network capacity and speed, each shard operates independently.
Disadvantages	Slower processing, higher costs, scalability issues, limited privacy.	Less secure, centralized intermediaries, trust dependency, interoperability challenges.	Relies on a common ledger, faces synchronization challenges between shards.

Table 1: *Comparison of On-chain vs Off-chain solution with sharding.[9][10][11]*

Zilliqa: Zilliqa's approach to sharding is different from Ethereum's. It implements sharding without a central coordinator. In Zilliqa's system, all single-shard transactions are executed in parallel but transactions affecting the same smart contract or more than one shard are not executed in parallel. This helps in increasing the processing power and distribution of information across the network still there are some limitations like the requirement for sApps to reside in most of the shards.[8]

4. Comparison of on-chain vs off-chain solutions:

The comparison of on-chain, off-chain transactions, and sharding in blockchain (refer **Table 1**) reveals different characteristics, advantages and disadvantages. On-chain transactions are processed directly on the blockchain they are known for their security, transparency, and support for smart contracts, but suffer from slower processing and higher costs due to the consensus mechanism. Off-chain transactions offer a faster more cost effective alternative with better privacy and flexibility, but they often involve centralized intermediaries and have security concerns. Sharding utilizes the on-chain mechanism to increase network capacity and speed by dividing the blockchain into independent segments. However, it still relies on a common ledger and faces synchronization and scalability challenges. This comparison highlights the balance between efficiency, security, and decentralization in blockchain technologies[9][10][11].

5. Future Scope:

The future scope of sharding in blockchain technology might be enhancing its scalability, security, and efficiency. The main areas of focus are:

5.1 Minimizing Global Chain Dependency:

Future research cn be done in developing decentralized methods for key operations like randomness generation which are important for shard integrity and security. This could be done by exploring new algorithms that reduce the need for a centralized global chain thesreby avoiding centralization and bottleneck issues.[12][14]

5.2 Optimizing Storage and Communication Overhead:

Future advancements might focus on improving data storage efficiency and communication protocols between shards. This could help in innovating data compression techniques and more efficient consensus protocols that reduce the burden on nodes.[13][14]

5.3 Rethinking Node Allocation:

Strategies could involving more dynamic and adaptive node allocation methods can be looked into to optimize shard sizes and compositions. This could mean developing algorithms that consider network load, transaction types, and node capabilities to balance workload and minimize intra-shard latency.

5.4 Improving Intra-Consensus Security:

Researchers can focus on developing more sophisticated algorithms for randomness generation including new security protocols that are specifically tailored for the unique challenges of sharded blockchain networks.[14]

5.5 Revamping Proof-of-Work (PoW) Consensus:

Efforts could be made to make PoW consensus more suitable for sharded environments, potentially by removing interdepencence of certain resource-intensive aspects of PoW and PoS from the primary transaction validation process.[14]

5.6 Facilitating Efficient Cross-Shard Transactions:

Ongoing research may focus on creating more efficient protocols for managing cross-shard transactions that are important for intensive operations like smart contracts. This could involve the development of new transaction protocols that reduce latency and overhead in cross-shard communication.[14]

6. Conclusion:

In conclusion, sharding presents a key advancement in blockchain technology, which addresses the issue of scalability. By dividing the network into smaller more manageable shards it significantly enhances transaction speed and network efficiency. The paper has explored various sharding techniques and their impact on scalability, along with real-world applications and challenges. Looking ahead, the continuous evolution in sharding strategies particularly in minimizing global chain dependency, optimizing storage and communication and improving consensus mechanisms, will be crucial. This research will not only enhance the performance of blockchain networks but also broaden their application scope making blockchain technology more adaptable and robust for diverse uses. As sharding matures, it is poised to play a key role in the future of decentralized systems marking a significant step towards solving blockchain's scalability dilemma.

7. References:

1. IBM. (n.d.). [Blockchain Overview]. Retrieved from https://www.ibm.com/topics/blockchain

2. Crypto.com University. (JUN 30, 2023). What is Sharding? Retrieved from https://crypto.com/university/what-is-sharding

3. Yi Li, Jinsong Wang, Hongwei Zhang, A survey of state-of-the-art sharding blockchains: Models, components, and attack surfaces, Journal of Network and Computer Applications, Volume 217, 2023, 103686, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2023.103686.

4. Yizhong Liu, Jianwei Liu, Marcos Antonio Vaz Salles, Zongyang Zhang, Tong Li, Bin Hu, Fritz Henglein, Rongxing Lu et al. Building blocks of sharding blockchain systems: Concepts, approaches, and open problems, Computer Science Review, Volume 46, 2022, 100513, ISSN 1574-0137, https://doi.org/10.1016/j.cosrev.2022.100513.

5. Huhnstock, R., Reginka, M., Sonntag, C. et al. Three-dimensional close-to-substrate trajectories of magnetic microparticles in dynamically changing magnetic field landscapes. Sci Rep 12, 20890 (2022). https://doi.org/10.1038/s41598-022-25391-z

6. Runchao Han, Jiangshan Yu, Haoyu Lin, Shiping Chen, Paulo Esteves-Veríssimo et al. On the Security and Performance of Blockchain Sharding, Cryptology ePrint Archive, Paper 2021/1276, https://eprint.iacr.org/2021/1276

7. De Jong, L. (2023, July 20). Sharding: The Key to Scaling Blockchain Networks. onXRP.com. Retrieved from https://onxrp.com/sharding-the-key-to-scaling-blockchain-networks/.

8. Coin Rivet. (n.d.). Four Projects Leading the Way in Database Sharding. Retrieved from https://coinrivet.com/guides/altcoins/four-projects-leading-the-way-in-database-sharding/

9. Abrol, A. (2022, September 7). Crypto Off-Chain vs On-Chain. Blockchain Council. Retrieved from https://www.blockchain-council.org/blockchain/crypto-off-chain-vs-on-chain/.

10. Price, A. (2023, May 2). On-Chain vs Off-Chain. Tectum. Retrieved from https://tectum.io/blog/on-chain-vs-off-chain/.

11. Gupta, A. (2019, August 13). Sharding vs Off-Chain: What's Better for Blockchain Scalability? CoinFunda. Retrieved from https://coinfunda.com/sharding-vs-off-chain-whats-better-for-blockchain-scalability/.

12. H. Chen and Y. Wang, "SSChain: A full sharding protocol for public blockchain without data migration overhead," Pervasive Mobile Comput., vol. 59, Oct. 2019, Art. no. 101055. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119218306370.

13. M. Al-Bassam, A. Sonnino, and V. Buterin, "Fraud and data availability proofs: Maximising light client security and scaling blockchains with dishonest majorities," CoRR, vol. abs/1809.09044, pp. 1–34, Sep. 2018. [Online]. Available: http://arxiv.org/abs/1809.09044

14. G. Yu, X. Wang, K. Yu, W. Ni, J. A. Zhang and R. P. Liu, "Survey: Sharding in Blockchains," in IEEE Access, vol. 8, pp. 14155-14181, 2020, doi: 10.1109/ACCESS.2020.2965147.

15. Jon, C., 2022. The Hitchhiker's guide to ethereum. (Online; Accessed 30 April 2023). https://members.delphidigital.io/reports/the-hitchhikers-guide-to-ethereum/

16. Abdurrashid Ibrahim Sanka, Ray C.C. Cheung, A systematic review of blockchain scalability: Issues, solutions, analysis and future research, Journal of Network and Computer Applications, Volume 195, 2021, 103232, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2021.103232.

17. Mearian, L. (2019, August 12). Why blockchain could be a threat to democracy. Computerworld. Retrieved from https://www.computerworld.com/article/3430697/why-blockchain-could-be-a-threat-to-democracy.html