sensors

MDPI

Article

Sharding-Based Proof-of-Stake Blockchain Protocols: Key Components & Probabilistic Security Analysis †

Abdelatif Hafid [1,*], Abdelhakim Senhaji Hafid [1] and Dimitrios Makrakis [2]

[1] Department of Computer Science and Operations Research, Université de Montréal, Montreal, QC H3T 1J4, Canada
[2] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada
* Correspondence: abdelatif.hafid@yahoo.com

† This paper is an extended version of our paper published in Hafid, A.; Hafid, A.S.; Senhaji, A. Sharding-Based Proof-of-Stake Blockchain Protocol: Security Analysis. In Proceedings of the 4th International Congress on Blockchain and Applications, Lecture Notes in Networks and Systems, L'Aquila, Italy, 13–15 July 2022; Springer: Berlin/Heidelberg, Germany, 2022.

Abstract: Blockchain technology has been gaining great interest from a variety of sectors including healthcare, supply chain, and cryptocurrencies. However, Blockchain suffers from a limited ability to scale (i.e., low throughput and high latency). Several solutions have been proposed to tackle this. In particular, sharding has proved to be one of the most promising solutions to Blockchain's scalability issue. Sharding can be divided into two major categories: (1) Sharding-based Proof-of-Work (PoW) Blockchain protocols, and (2) Sharding-based Proof-of-Stake (PoS) Blockchain protocols. The two categories achieve good performances (i.e., good throughput with a reasonable latency), but raise security issues. This article focuses on the second category. In this paper, we start by introducing the key components of sharding-based PoS Blockchain protocols. We then briefly introduce two consensus mechanisms, namely PoS and practical Byzantine Fault Tolerance (pBFT), and discuss their use and limitations in the context of sharding-based Blockchain protocols. Next, we provide a probabilistic model to analyze the security of these protocols. More specifically, we compute the probability of committing a faulty block and measure the security by computing the number of years to fail. We achieve a number of years to fail of approximately 4000 in a network of 4000 nodes, 10 shards, and a shard resiliency of 33%.

Keywords: security analysis; blockchain; probabilistic analysis; sharding-based blockchain protocols; malicious nodes; proof of stake; practical Byzantine fault tolerance

Citation: Hafid, A.; Hafid, A.S.; Makrakis, D. Sharding-Based Proof-of-Stake Blockchain Protocols: Key Components & Probabilistic Security Analysis. *Sensors* **2023**, *23*, 2819. https://doi.org/10.3390/s23052819

Academic Editors: Javier Prieto and Ramón J. Durán Barroso

Received: 5 December 2022
Revised: 16 February 2023
Accepted: 27 February 2023
Published: 4 March 2023

1. Introduction

With the rise of Bitcoin [1], Blockchain has attracted significant attention from both industry and academia. More specifically, it has been adopted across different industries including healthcare [2,3], finance [4], and the public sector [5]. However, Blockchain suffers from poor scalability [6]. For example, in the case of cryptocurrencies, Bitcoin [1] handles between three and seven transactions per second (tx/s), which is very limited compared to traditional electronic payment systems (e.g., PayPal [7]). Several solutions were proposed to scale Blockchain. In particular, sharding has emerged as a promising solution [6]. Sharding consists of partitioning the network into sub-networks, called shards. All shards work in parallel to enhance the performance of the network. More specifically, each shard processes a sub-set of transactions instead of the entire network processing all the transactions. While sharding considerably improves scalability, it decreases the level of Blockchain security. More specifically, in sharding-based Blockchains, it is easier for a malicious user to attack and conquer a single shard compared to the whole network. This attack is well-known as a shard takeover attack (also referred to as a 1% attack) [8].

Blockchain networks are susceptible to sybil attacks by malicious nodes (called sybil nodes). Several consensus mechanisms (e.g., PoW, PoS, and pBFT) have been proposed to defend against these sybil nodes. Sharding-based Blockchain protocols [6,9] can be classified into two classes: sharding-based PoW and sharding-based PoS Blockchain protocols.

Recently, Hafid et al. [8,10,11] proposed mathematical models to analyze the security of sharding-based PoW protocols. However, to the best of our knowledge, there is no existing work that proposes a probabilistic security analysis of sharding-based PoS Blockchain protocols except an earlier version of this paper, which has been published in [12].

In this paper, we focus on the second category. We start by presenting the key components of these protocols. This article briefly describes PoS and pBFT consensus mechanisms, and discusses their use and limitations in the context of sharding-based Blockchain protocols. Finally, we propose a probabilistic model to analyze the security of these protocols by computing the probability of committing a faulty block. Based on these probabilities, we calculate the number of years to fail for the purpose of quantifying and measuring the security of the network.

The remainder of the paper is organized as follows. Section 2 presents the key components of sharding-based PoS Blockchain protocols and compares these protocols with related Blockchain protocols. Section 3 presents the proposed probabilistic model. Section 4 presents the numerical results and evaluates the proposed model. Section 5 present limitations of the paper and future work. Lastly, concluding remarks are given in Section 6.

2. Key Components of Sharding-Based PoS Blockchain Protocols

In this section, we cover the main components of shading-based PoS Blockchain protocols and compare these protocols with Bitcoin and sharding-based PoW Blockchain protocols.

2.1. Key Components

In this section, we shed light on the key components of sharding-based PoS Blockchain protocols. More specifically, we focus on Incognito [13] as an example. However, most of the sharding-based PoS Blockchain protocols share the same structure (see Figure 1). The key components of these protocols are:

- **Validator:** A node (can be malicious or honest) that competes to produce and add blocks to the blockchain.
- **Consensus mechanism:** Most of the sharding-based PoS Blockchain protocols use pPFT alongside PoS. PoS is an alternative consensus mechanism to PoW. Unlike PoW, which relies on miners with hash power to solve a mathematical puzzle, PoS relies on validators staking coins. Validators are selected based on the number of coins they stake (e.g., a validator with 10% of the total coins staked, has a 10% probability of adding a block to the blockchain). PoS consists of selecting validators in proportion to their number of coins. Validators are responsible for adding new blocks to Blockchain. pBFT is an algorithm that tolerates Byzantine faults [14]; An algorithm belonging to the Byzantine Fault Tolerance (BFT) class. BFT is the ability of the network to reach a consensus, on some value, even in the case of having some failed and/or malicious nodes in the network. Lamport et al. [15] proved that if we have $3m + 1$ correctly working processors, a consensus (agreement on same value) can be reached if at most m processors are faulty; this means that strictly more than two-thirds of the total number of processors should be honest.
- **Beacon Chain:** The main chain in the network. It is responsible for randomly assigning the validators to the committees of shards. The Beacon chain confirms cross-shard information and shuffles the committees of shards to ensure security.
- **Chain (aka shard):** A chain that consists of a subset of nodes of the network. Shards process different transactions in parallel to improve scalability. Generally, a sharding-based PoS network consists of one beacon chain and several shards.

- **Beacon Committee:** A subset of validators, selected randomly from the set of validators, that has decided to stake for the beacon chain. Generally, its main role is to check/verify and insert the valid block header into the beacon chain.
- **Shard Committee:** A subset of validators, selected randomly from the set of validators, that has decided to stake for the shard chain. A shard committee validates and confirms the transactions processed in the shard.

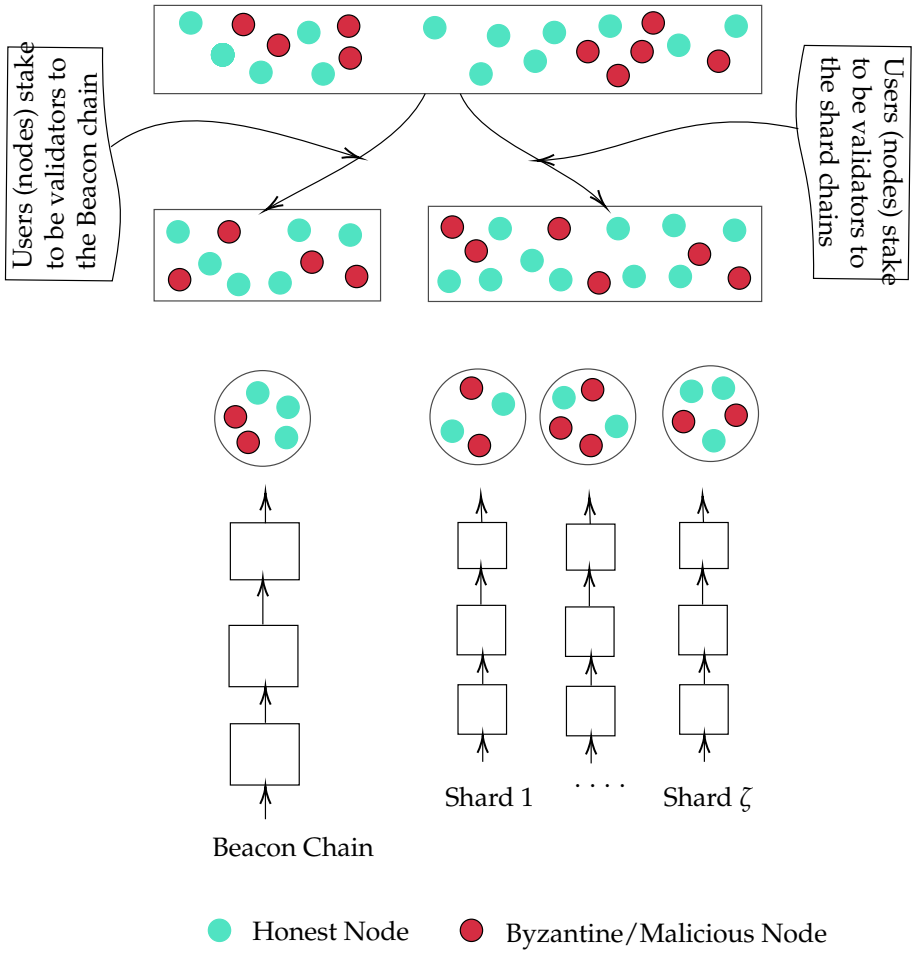


Figure 1. A sharding-based PoS and pBFT Blockchain protocol.

2.2. Comparison with Related Blockchain Protocols

To demonstrate the advantages and limitations of sharding-based PoS Blockchain protocols, Table 1 shows a comparison of its performance with sharding-based PoW Blockchain protocols and PoS Blockchain protocols. This comparison reflects that sharding-based PoS Blockchain protocols have several good features. Specifically, sharding-based PoS Blockchain protocols have higher transaction throughput, requires less computational power, and in some cases offers higher level of privacy (e.g., Incognito [13]) and lower transactions fees. However, some limitations still remain in sharding-based PoS Blockchain protocols; for example, centralization concerns caused by coinage.

Table 1 shows that by using sharding alongside PoS and pBFT consensus mechanisms, we get better performance without the need for high hash computing power. For example, the throughput of Incognito [13] scales out linearly with the number of shards. Specifically, Incognito achieves a significantly higher number of transactions (up to 800 tx/s for only 64 shards; thus, for 100 s of shards, the throughput will be in 1000 s tx/s). Zilliqa [16] handles up to 2800 tx/s [17].

Table 1. Bitcoin vs. Sharding-Based PoW Blockchain Protocols vs. Sharding-Based PoS Blockchain Protocols.

	Bitcoin [1]	**Zilliqa [16]**	**Elastico [18]**	**Incognito [13]**	**Harmony [17]**	**Nxt [19]**
PoW	✓	✓	✓	×	×	×
PoS	×	×	×	✓	✓	✓ [1]
pBFT	×	✓	✓	✓	✓	×
Sharding	×	✓	✓	✓	✓	×
Throughput	Up to 7 tx/s	2800 tx/s	40 tx/s	100 [1] tx/s and 800 [2] tx/s	N/A	4 tx/s
Latency	∼1 h	N/A	800 s	N/A	N/A	∼10 min
Resiliency	Supports up to 50% ($\lesssim$51%) of Byzantine fault	Up to 33% for shard's committee and 25% for the entire network	Up to 33% for shard's committee and 25% for the entire network	Up to 33% for shard's committee and 51% for beacon's committee	Up to 33% for shard's committee and 25% for the entire network	Supports up to 33% ($\lesssim \frac{1}{3}$) of Byzantine fault participants
Unique Features	Mining competition	Proposes an innovative smart contract language	First sharding protocol with presence of Byzantine fault	BLS [a] consensus	State sharding (i.e., Harmony shards Blockchain state)	The miner of the next block is predictable
Drawbacks	High computational power and low performance	transactions sharding [b]	Uses small committee size	Centralization concern due to coinage	Centralization concern due to coinage	Centralization concern due to coinage
Advantages	High level of security	Uses PoW as identity registration to prevent Sybil attacks	Ensures good randomness	High level of privacy	Provides consistent cross-shard transactions	Agile architecture

✓ : has property; ×: does not have property; N/A: **N**ot **A**vailable in the literature. [1] 8 shards; [2] 64 shards . [a] beside PoS and pBFT, Incognito [13] use an additional consensus, BLS; Incognito implements BLS for multi-signature aggregation; [b] each node has to hold the entire Blockchain state to be able to process transactions.

Incognito [13] claims significantly higher performance compared to other privacy Blockchains (e.g., Nxt [19] and Zcash [20]). Nxt [19] can usually handle less than 10 tx/s while Zcash handles only 6 tx/s.

3. Probabilistic Model

In this section, we propose a probabilistic model to analyze the security of a sharding-based PoS Blockchain protocols, called Incognito [13].

3.1. Notations & Architecture

This section describes the notations and definitions used to represent the proposed probabilistic model as well as the architecture of Incognito [13].

3.1.1. Notations & Definitions

Table 2 shows the notations and symbols that are used throughout the paper.

Definition 1 (Faulty block)**.** *A faulty block is a block that contains fraudulent transactions.*

Definition 2 (Conquering the Protocol)**.** *A protocol is said to be conquered if the malicious nodes succeed in adding a faulty block to the blockchain.*

Definition 3 (Committee Resiliency of a Shard)**.** *The maximum percentage of malicious nodes that the committee of the shard chain can contain while remaining secure.*

Definition 4 (Committee Resiliency of the Beacon Chain)**.** *The maximum percentage of malicious nodes that the committee of the beacon chain can support while remaining secure.*

Table 2. Notations & Symbols.

Notation	Description
$\mathcal{N}$	Number of users (aka number of nodes)
n	Committee size of a shard
n'	Committee size of the beacon chain
H	Number of honest validators in a shard
M	Number of malicious validators in a shard
$\mathcal{V}$	Number of validators in a shard ($\mathcal{V} = H + M$)
ζ	Number of shards
X	Random variable that computes the number of malicious nodes in the committee of a shard
H'	Number of honest validators in the beacon chain
M'	Number of malicious validators in the beacon chain
$\mathcal{V}'$	Number of validators in the beacon chain ($\mathcal{V}' = H' + M'$)
X'	Random variable that computes the number of malicious nodes in the committee of the beacon chain
r	Resiliency of the shard committee, $0 \leq r \leq 1$
r'	Resiliency of the beacon committee, $0 \leq r' \leq 1$
R	Percentage of malicious validators in a shard chain
R'	Percentage of malicious validators in the beacon chain
$\mathcal{P}_f$	Probability of conquering the protocol
$\mathcal{P}$	Probability of a shard committing a faulty block
$\mathcal{P}^{'}$	Probability of all shards committing a faulty block
$\mathcal{P}^{''}$	Probability of the beacon chain committing a faulty block
$\mathcal{Y}_f$	Number of years to fail

3.1.2. Architecture of Incognito

Figure 1 shows the structure/architecture of Incognito [13], a sharding-based PoS Blockchain protocol. The network contains a single beacon chain and ζ shard chains. Each user/node can stake to be a validator either for the beacon chain or for the shard chain (see Figure 1). Shard chains produce blocks in parallel. All shard chains are synchronized by the beacon chain. More specifically, each shard has its own committee, which is randomly assigned by the beacon chain. Each shard chain processes a subset of the transactions submitted to the network. When a shard block is created, the beacon committee verifies the block; if it is valid, it adds the block header to the beacon chain. Otherwise, it drops it and sends the proof to other shards for a vote to slash the misbehaving shard committee. Furthermore, in each epoch, the beacon chain shuffles the committees of the shards to ensure security. For Incognito [13], when a new random number is generated, the beacon chain shuffles the committees; one epoch, for Incognito, corresponds to generating a new random number. This number is generated periodically in a round-robin fashion [13,21].

3.2. Probability Distributions

The main idea behind the sharding solution is to split/divide the network into subsets, called shards. Every single shard processes a subset of transactions rather than the entire network processing all transactions. This idea allows the network to scale (in terms of the number of transactions per second) with the number of shards. However, this technique may compromise the security of the network [8].

In Incognito [13], to add a faulty block to the Blockchain, it must be confirmed by at least $\lfloor \beta n \rfloor$ ($0 \leq \beta \leq 1$; $\beta = r$) of the shard committee members, by at least $\lfloor \gamma n' \rfloor$ of the beacon committee members ($\gamma = r'$), and by at least $\lfloor \delta \zeta \rfloor$ ($0 \leq \delta \leq 1$) of all shards' committees. For *Incognito* [13], $\beta = \delta = \frac{2}{3}$ and $\gamma = \frac{1}{2}$.

Let X be a random variable that represents the number of malicious nodes sampled in the committee of a shard from the validators in that shard. Each of the sampled nodes can be placed in one of 2 distinctive and disjoint groups; honest nodes group or malicious nodes group. Because the committees do not overlap, we have sampling without replacement. Thus, we conclude that X follows the hypergeometric distribution with parameters $\mathcal{V}$, M, and n (it can be written as follows: $X \sim \mathbf{H}(\mathcal{V}, M, n)$).

The probability mass function corresponds to X can be defined as follows [22]:

$$P(X = \omega) = \frac{\binom{M}{\omega}\binom{\mathcal{V}-M}{n-\omega}}{\binom{\mathcal{V}}{n}} \tag{1}$$

where $\max(0, n - \mathcal{V}) \leq \omega \leq \min(M, n)$ and $\mathcal{V} - M = H$.

Lemma 1. *The probability of a shard's committee to commit a faulty block ($\mathcal{P}$) can be expressed as follows:*

$$P(X \geq \lfloor \beta n \rfloor) = \sum_{i=\lfloor \beta n \rfloor}^{n} \frac{\binom{M}{i}\binom{\mathcal{V}-M}{n-i}}{\binom{\mathcal{V}}{n}} \tag{2}$$

Proof of Lemma 1 results directly from the cumulative hypergeometric distribution [8,11].

Similarly to X, X' follows the hypergeometric distribution with parameters $\mathcal{V}'$, M', and n'.

Lemma 2. *The probability of at least $\lfloor \delta\zeta \rfloor$ shards committees committing a faulty block ($\mathcal{P}'$) can be computed as follows:*

$$\sum_{k=\lfloor \delta\zeta \rfloor}^{\zeta} \left(P(X \geq \lfloor \beta n \rfloor) \right)^k = \sum_{k=\lfloor \delta\zeta \rfloor}^{\zeta} \left[\sum_{i=\lfloor \beta n \rfloor}^{n} \frac{\binom{M}{i}\binom{\mathcal{V}-M}{n-i}}{\binom{\mathcal{V}}{n}} \right]^k \tag{3}$$

Proof of Lemma 2. The minimum number of committees to commit a faulty block is $\lfloor \delta\zeta \rfloor$, where ζ is the number of shards. The probability of exactly $\lfloor \delta\zeta \rfloor$ committees confirm/agree to add a faulty block can be expressed as follows:

$$P_{\lfloor \delta\zeta \rfloor} = \left(P(X \geq \lfloor \beta n \rfloor) \right)^{\lfloor \delta\zeta \rfloor} \tag{4}$$

The probability to commit a faulty block by exactly $\lfloor \delta\zeta \rfloor + 1$ committees can be expressed as follows:

$$P_{\lfloor \delta\zeta \rfloor} + 1 = \left(P(X \geq \lfloor \beta n \rfloor) \right)^{\lfloor \delta\zeta \rfloor + 1} \tag{5}$$

Similarly, the probability of exactly ζ committees (the entire number of shards in this case) agreeing to add a faulty block can be expressed as follows:

$$P_{\zeta} = \left(P(X \geq \lfloor \beta n \rfloor) \right)^{\zeta} \tag{6}$$

A faulty block can be committed if $\lfloor \delta\zeta \rfloor$ or $\lfloor \delta\zeta \rfloor + 1$ or $\lfloor \delta\zeta \rfloor + 2$, $\cdots$, or ζ committees agree to add this block. This can be mathematically computed by the sum over all these probabilities and can be expressed as follows:

$$\mathcal{P}'' = P_{\lfloor \delta\zeta \rfloor} + P_{\lfloor \delta\zeta \rfloor + 1} + \cdots + P_{\zeta} \tag{7}$$

$$= \sum_{i=0}^{\zeta - \lfloor \delta\zeta \rfloor} P_{\lfloor \delta\zeta \rfloor + i} \tag{8}$$

□

Lemma 3. *The probability of the beacon's committee committing a faulty block ($\mathcal{P}''$) can be expressed as follows:*

$$P(X' \geq \lfloor \gamma n' \rfloor) = \sum_{j=\lfloor \gamma n' \rfloor}^{n'} \frac{\binom{M'}{j}\binom{\mathcal{V}'-M'}{n'-j}}{\binom{\mathcal{V}'}{n'}} \tag{9}$$

where $\mathcal{V}' - M' = H'$.

Proof of Lemma 3 results directly from the cumulative hypergeometric distribution [8,11].

Theorem 1 (Committing a Faulty Block)**.** *The probability of committing a faulty block ($\mathcal{P}_f$) by a given sharding-based PoS Blockchain protocol can be expressed as follows:*

$$\mathcal{P}_f = \sum_{k=\lfloor \beta n \rfloor}^{n} \sum_{j=\lfloor \gamma n' \rfloor}^{n'} \sum_{\alpha=\lfloor \delta \zeta \rfloor}^{\zeta} \left[\frac{\binom{M}{k}\binom{H}{n-k}\binom{M'}{j}\binom{H'}{n'-j}}{\binom{\mathcal{V}}{n}\binom{\mathcal{V}'}{n'}} \left(\sum_{i=\lfloor \beta n \rfloor}^{n} \frac{\binom{M}{i}\binom{\mathcal{V}-M}{n-i}}{\binom{\mathcal{V}}{n}} \right)^{\alpha} \right] \tag{10}$$

Proof of Theorem 1. To commit a faulty block, it must be confirmed/verified by at least $\lfloor \beta n \rfloor$ of the shard committee members, by at least $\lfloor \gamma n' \rfloor$ of the beacon committee members, and by at least $\lfloor \delta \zeta \rfloor$ of all shards' committees. This can be expressed by the product over the three probabilities (the calculated probabilities in Lemmas 1, 2, and 3). □

3.3. Years to Fail

To make the measurement of the security more readable, we propose to compute the average number of years to fail ($\mathcal{Y}_f$) based on the calculated failure probability (i.e., the probability of conquering the protocol). This number can be expressed as follows:

$$\mathcal{Y}_f = \frac{1}{\mathcal{P}_f \mathcal{N}_s} \tag{11}$$

where $\mathcal{P}_f$ is the probability of committing (adding) a faulty block to the blockchain and $\mathcal{N}_s$ is the number of sharding rounds per year (also referred to as the number of epochs per year).

4. Evaluation Results

In this section, we evaluate the effectiveness of the proposed probabilistic model via numerical simulations.

4.1. Simulation Setup

In order to implement the proposed probabilistic model, we make use of a built-in Python library called **SciPy**. We import **hypergeom**, a hypergeometric discrete random variable, from the **scipy.stats** module. Specifically, we use the **cdf** (i.e., cumulative distribution function) function. We also make use of the **math** module, which provides access to mathematical functions (e.g., the **floor** function). For the results plot, we make use of **matplotlib.pyplot** library.

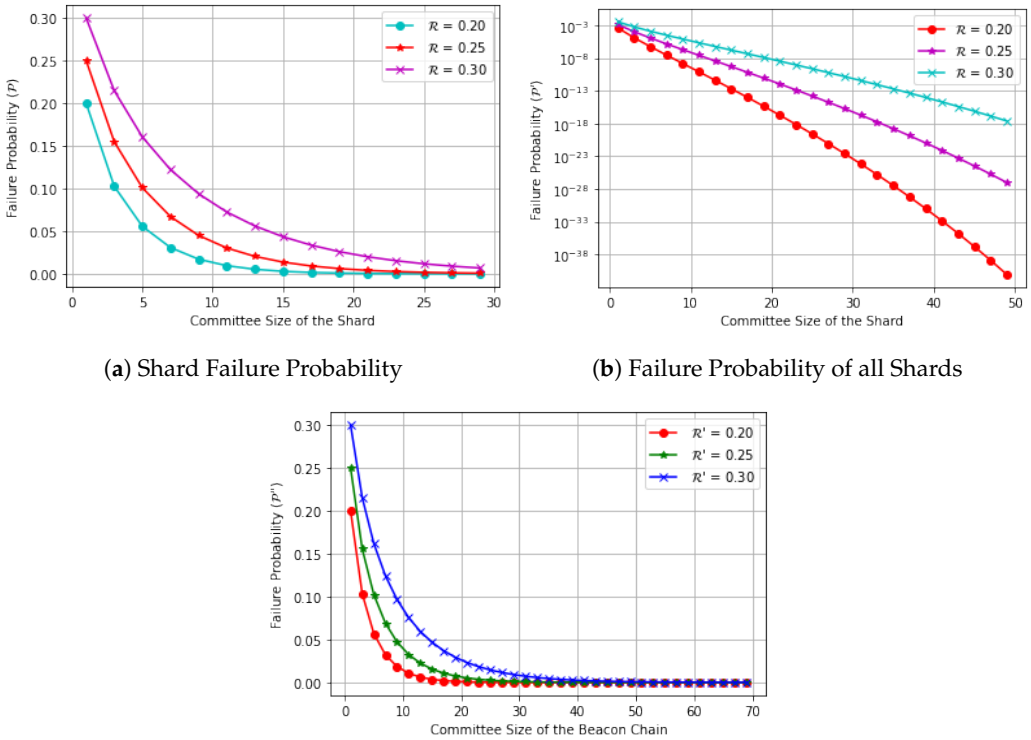
4.2. Results and Analysis

In Figure 2, we assume a network with $\mathcal{N}$= 2000 nodes, $\mathcal{V}$ = 200, $\mathcal{V}'$ = 400, ζ = 8, $r = r'$ = 0.5. We chose these values for r and r' to make the results more readable. However, we can use other values (e.g., Incognito [13] values). Figure 2 illustrates how the committee size of the shard impacts its failure probability $\mathcal{P}$ as well as the failure probability of all shards $\mathcal{P}'$ and the relationship between the committee size of the beacon chain and its failure probability $\mathcal{P}''$. In particular, Figure 2a shows the probability of a shard to commit a faulty block versus the size of the committee. We observe that the probability $\mathcal{P}$ decreases

when the size of the committee increases. More specifically, we observe that the probability corresponding to $\mathcal{R}$ = 0.2 (i.e., 20% of malicious nodes in each shard) decreases rapidly compared to those of $\mathcal{R}$ = 0.25 and $\mathcal{R}$ = 0.3; this can be explained by the small percentage of malicious nodes. In other words as the percentage of malicious nodes decreases, so does the probability.

Figure 2b shows the probability of all shards committing a faulty block versus the size of the committee. We observe that the probability $\mathcal{P}^{'}$ decreases when the size of the committee increases. Similarly, as the percentage of malicious nodes slightly increases in the shard, the probability of committing a faulty block increases.

Figure 2c shows the probability of the beacon chain to commit a faulty block ($\mathcal{P}''$) versus the size committee of the beacon chain (n'). We also observe that the probability $\mathcal{P}^{''}$ decreases when the size of the committee increases. More specifically, we observe that the probability corresponding to $\mathcal{R} = 0.2$ (i.e., 20% of malicious nodes in the beacon chain) decreases sharply compared to those of $\mathcal{R} = 0.25$ and $\mathcal{R} = 0.3$.



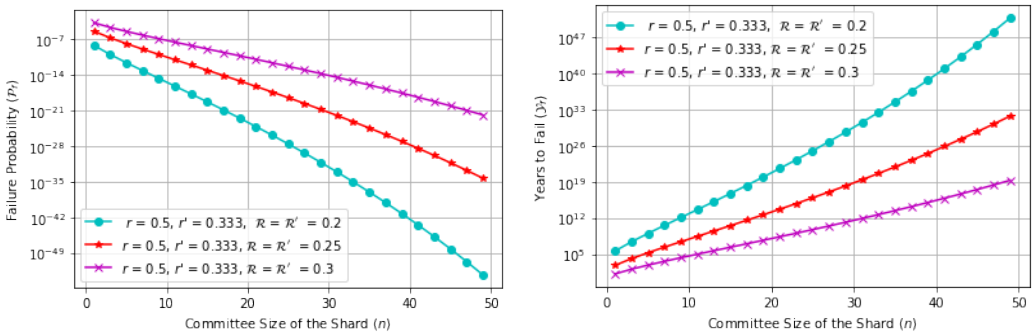
(**a**) Shard Failure Probability

(**b**) Failure Probability of all Shards

(**c**) Beacon Failure Probability

Figure 2. (**a**) Probability of a shard to commit a faulty block ($\mathcal{P}$) versus the committee size of the shard (n), (**b**) Log-scale plot of the probability of all shards committing a faulty block ($\mathcal{P}^{'}$) versus the size of the committee (n), and (**c**) Probability of the beacon chain to commit a faulty block ($\mathcal{P}''$) versus the size committee of the beacon chain (n').

In Figure 3, we assume a network with 2000 nodes, $\mathcal{V} = 200$, $\mathcal{V}' = 400$, $\zeta = 8$, $n' = 100$. Figure 3a shows the probability of conquering the protocol when varying the committee size of the shard. We observe that as the committee size of the shard increases the probability of conquering the protocol decreases. Figure 3b shows the number of years to fail ($\mathcal{Y}_f$) versus the committee size of the shard. We observe that when the committee size of the shard increases, the number of years to fail increases as well.



(**a**) Probability of Conquering the Protocol (**b**) Years to Fail

Figure 3. Log-scale plot: (**a**) Probability of conquering the protocol ($\mathcal{P}_f$) versus the committee size of the shard (n), (**b**) Number of years to fail ($\mathcal{Y}_f$) versus the committee size of the shard (n).

Table 3 shows the probability of conquering the chain (i.e., the probability of committing a faulty block; it is calculated based on Theorem 1) for different percentages of malicious nodes in the shards as well as in the beacon chain. Moreover, Table 3 shows the number of years to fail corresponding to these probabilities. We observe that as the percentage of malicious nodes increases the number of years to fail decreases. More specifically, we observe that the probability of conquering the chain is extremely low even with 20% of the malicious nodes in each shard as well as in the beacon chain. This achieves good security, reaching about $1.74E + 17$ years to fail.

Table 3. Probability of conquering the protocol.

$\mathcal{R} = \mathcal{R}'$	10%	15%	20%	30%
$\mathcal{P}_f$ [1]	3.63E-66	2.10E-34	1.58E-18	1.70E-04
$\mathcal{Y}_f$ [1]	7.56E+62	1.30E+31	1.74E+17	16.12
$\mathcal{P}_f$ [2]	0.0	5.14E-80	2.01E-41	5.30E-07
$\mathcal{Y}_f$ [2]	inf	5.33E+76	1.36E+38	5171.32

[1] Scenario 1. [2] Scenario 2.

Table 4 shows the trade-off between the values of n and n' that provide certain target value of years to fail. We consider a network with 2000 nodes, $\mathcal{V} = 200$, $\mathcal{V}' = 400$, $\zeta = 8$, $\mathcal{R} = \mathcal{R}' = 0.3$, and $\delta = r = r' = 0.33$. To reach a level of security corresponding to 2500 years to fail, you should set n to 95 and n' to 150 or n to 85 and n' to 145. However, you should adjust the values of n and n' carefully to reach the desired target. Similarly, if you have a network with 4000 nodes, $\mathcal{V} = 300$, $\mathcal{V}' = 1000$, $\zeta = 10$, $\mathcal{R} = \mathcal{R}' = 0.3$, and $\delta = r = r' = 0.33$ and set years to fail to as 4000, we can achieve it by setting n to 100 and n' to 250 or n to 122 and n' to 130. It is evident that there are different possibilities and combinations of the values of n and n'. However, you should adjust these values (n and n') carefully to reach the desired target (the desired number of years to fail ($\mathcal{Y}_f$)).

Table 4. Potential realistic scenarios.

$\mathcal{N}$	$\mathcal{V}$	$\mathcal{V}'$	ζ	$\mathcal{R}$	$\mathcal{R}'$	δ	r	r'	n	n'	Target ($\mathcal{Y}_f$)
2000	200	400	8	0.3 (30%)	0.3 (30%)	0.33 (33%)	0.33 (33%)	0.33 (33%)	95	150	≈2500
2000	200	400	8	0.3 (30%)	0.3 (30%)	0.33 (33%)	0.33 (33%)	0.33 (33%)	85	145	≈2500
4000	300	1000	10	0.3 (30%)	0.3 (30%)	0.33 (33%)	0.33 (33%)	0.33 (33%)	100	250	≈4000
4000	300	1000	10	0.3 (30%)	0.3 (30%)	0.33 (33%)	0.33 (33%)	0.33 (33%)	122	130	≈4000

Finally, we conclude that by adjusting the committee size of the shard as well as the committee size of the beacon chain, we can protect sharded Blockchain systems (based on PoS) against malicious nodes (e.g., Sybil nodes).

5. Limitations of the Paper and Future Work

Our study has some limitations, including the fact that the proposed probabilistic model is similar to that of Incognito [13]. However, each sharding-based PoS Blockchain protocol bears a specific structure, which is slightly different from other sharding-based PoS Blockchain protocols (e.g., Harmony [17]). Furthermore, this study does not provide an in-depth literature review of sharding-based PoS Blockchain protocols since the objective was to provide a brief and simple presentation of the key components of these protocols. The aim of doing so was to prepare the readers to understand how we can deal with the security analysis of these protocols using probability distributions.

Future works will focus on analyzing the security of novel types of sharding-based Blockchain protocols including DankSharding by Ethereum [23], Shardus [24] and probabilistic analysis of cross-shard transactions.

6. Conclusions

In this paper, we first illustrate the key components of sharding-based PoS Blockchain protocols. Next, we discuss two consensus mechanisms, PoS and pBFT. We also address the security of Incognito, a sharding-based PoS Blockchain protocol. In particular, we provide a probabilistic model to compute the probability of committing a faulty block. Based on this probability, we compute the number of years to fail. Furthermore, this article depicts that we can control the number of years to fail by adjusting the size of the shard as well as the size of the beacon committee. Our future work includes the computation of the failure probability across shard transactions.

Author Contributions: Conceptualization, A.H.; Methodology, A.H.; Software, A.H.; Validation, D.M.; Writing—original draft, A.H.; Writing—review & editing, D.M.; Supervision, A.S.H. and D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nakamoto, S. Bitcoin Whitepaper. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 1 January 2022).
2. Kassab, M.H.; DeFranco, J.; Malas, T.; Laplante, P.; Destefanis, G.; Graciano Neto, V.V. Exploring Research in Blockchain for Healthcare and a Roadmap for the Future. *IEEE Trans. Emerg. Top. Comput.* **2019**, *9*, 1835–1852. [CrossRef]
3. Almaiah, M.A.; Hajjej, F.; Ali, A.; Pasha, M.F.; Almomani, O. A Novel Hybrid Trustworthy Decentralized Authentication and Data Preservation Model for Digital Healthcare IoT Based CPS. *Sensors* **2022**, *22*, 1448. [CrossRef] [PubMed]
4. Nordgren, A.; Weckström, E.; Martikainen, M.; Lehner, O.M. Blockchain in the fields of finance and accounting: A disruptive technology or an overhyped phenomenon. *ACRN Oxf. J. Financ. Risk Perspect.* **2019**, *8*, 47–58.
5. Abou Jaoude, J.; George Saade, R. Blockchain Applications—Usage in Different Domains. *IEEE Access* **2019**, *7*, 45360–45381. [CrossRef]
6. Hafid, A.; Hafid, A.S.; Samih, M. Scaling Blockchains: A Comprehensive Survey. *IEEE Access* **2020**, *8*, 125244–125262. [CrossRef]
7. PayPal. Available online: https://www.paypal.com/ca/home?locale.x=en_CA (accessed on 12 July 2022).
8. Hafid, A.; Hafid, A.S.; Samih, M. A tractable probabilistic approach to analyze sybil attacks in sharding-based blockchain protocols. *IEEE Trans. Emerg. Top. Comput.* 2022, *early access*. [CrossRef]
9. Wang, G.; Shi, Z.J.; Nixon, M.; Han, S. Sok: Sharding on blockchain. In Proceedings of the 1st ACM Conference on Advances in Financial Technologies, Zurich, Switzerland, 21–23 October 2019; pp. 41–61.
10. Hafid, A.; Hafid, A.S.; Samih, M. New Mathematical Model to Analyze Security of Sharding-Based Blockchain Protocols. *IEEE Access* **2019**, *7*, 185447–185457. [CrossRef]
11. Hafid, A.; Hafid, A.S.; Samih, M. A Novel Methodology-Based Joint Hypergeometric Distribution to Analyze the Security of Sharded Blockchains. *IEEE Access* **2020**, *8*, 179389–179399. [CrossRef]

12. Hafid, A.; Hafid Senhaji, A.; Senhaji, A. Sharding-Based Proof-of-Stake Blockchain Protocol: Security Analysis. In Proceedings of the 4th International Congress on Blockchain and Applications, L'Aquila, Italy, 13–15 July 2022; Springer: Berlin/Heidelberg, Germany, 2022.
13. Incognito. Available online: https://we.incognito.org/t/scaling-blockchain-privacy-with-dynamic-sharding/169 (accessed on 3 September 2022).
14. Castro, M.; Liskov, B. Practical byzantine fault tolerance. In *OSDI*; ACM: New Orleans, LA, USA, 1999; Volume 99, pp. 173–186.
15. Lamport, L.; Shostak, R.; Pease, M. The Byzantine generals problem. In *Concurrency: The Works of Leslie Lamport*; ACM: New York, NY, USA, 2019; pp. 203–226.
16. Zilliqa. The ZILLIQA Technical Whitepaper. Available online: https://docs.zilliqa.com/whitepaper.pdf (accessed on 10 August 2017).
17. Harmony. Technical Whitepaper. Available online: https://harmony.one/whitepaper.pdf (accessed on 10 July 2022).
18. Luu, L.; Narayanan, V.; Zheng, C.; Baweja, K.; Gilbert, S.; Saxena, P. A secure sharding protocol for open blockchains. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 17–30.
19. Nxt. Nxt: A Peer-to-Peer Digital Socioeconomic System. Available online: https://nxtdocs.jelurida.com/Nxt_Whitepaper (accessed on 15 March 2022).
20. Sasson, E.B.; Chiesa, A.; Garman, C.; Green, M.; Miers, I.; Tromer, E.; Virza, M. Zerocash: Decentralized anonymous payments from bitcoin. In Proceedings of the 2014 IEEE Symposium on Security and Privacy, Berkeley, CA, USA, 18–21 May 2014; pp. 459–474.
21. Rasmussen, R.V.; Trick, M.A. Round robin scheduling—A survey. *Eur. J. Oper. Res.* **2008**, *188*, 617–636. [CrossRef]
22. Rice, J.A. *Mathematical Statistics and Data Analysis*; Cengage Learning: Belmont, CA, USA, 2006.
23. Rafael, F. DankSharding. Available online: https://www.rootstrap.com/blog/danksharding-what-is-it-and-how-does-it-work (accessed on 2 December 2022).
24. Shardus. White Paper. Available online: https://shardus.com/whitepaper.pdf (accessed on 20 November 2021).