SASEUL: Universal Computer

Jungwoo Lee

jjal@artifriends.com

https://artifriends.com

Abstract. Satoshi Nakamoto proposed a way to select only one of the various databases on the network, which proved to be Bitcoin. This way was later named as a technology called blockchain. Vitalik Buterin introduced the concept of smart contracts in which data operates as a procedure for the first time in the blockchain, suggesting a way for the blockchain data itself to operate as a program, proving it as Ethereum. Through the blockchain, multiple computers can share the same data, code, and policy. In addition, it is possible to operate these with the same intention. However, in order to be introduced as a universal technology of mankind, there are three issues that need to be solved in performance. First, the cycle in which new data is uploaded is too long, secondly, it is difficult to store when the total data capacity increases because everyone must have the same data, and finally, the incentives for network participation are concentrated on a small number of people. This article deals with how to solve the above three problems.

1. Introduction

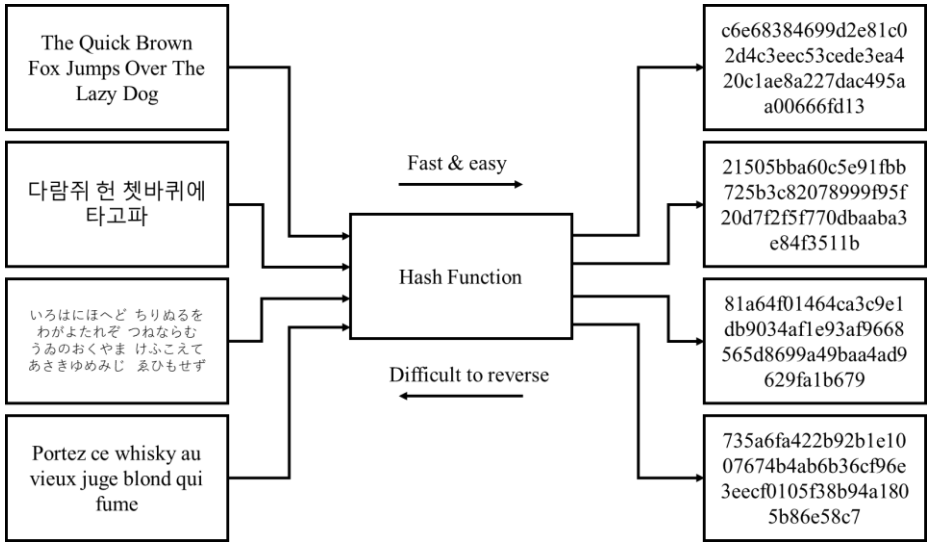
Modern computer systems have the problem that database administrators can manipulate data at any time. Satoshi Nakamoto proposed Bitcoin, an electronic cash system that operates only on Peer-to-Peer without a “trusted third party” to solve the problem.

In general, a trusted third party acts to unify the database of service users, and if there is no trusted third party to unify the database into one, a double-spending problem can occur. As a solution, Satoshi Nakamoto proposed the concept of blockchain and proof-of-work described as “The network timestamps transactions by hashing the chain of hash-based proof-of-work, forming a record that cannot be changed without re-doing the proof-of-work.” The principle of blockchain is as follows.

1.1. Hash function

The hash function refers to an irreversible function that converts an input value into a fixed length value that cannot be inversely transformed. Since the result value is a fixed length, there are more than one original input value with the same hash result value according to the pigeonhole principle.

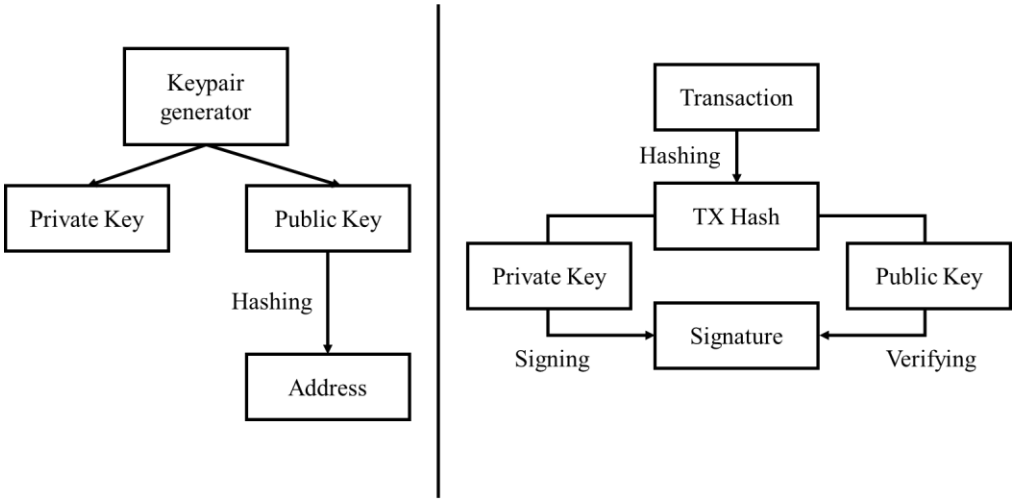
Although the hash result does not guarantee that it is calculated from any original input, it provides a clue that the result value is calculated from the corresponding original input with a high probability. In addition, it is difficult to infer the original input with only the hash result value, but knowing the original input value can verify that the hash result value was calculated from the corresponding original input value.



In the blockchain, the hash function is used to verify the integrity of the data. In a blockchain network, all systems must have the same data, and the same result must be derived when the same transaction is applied to the same data. In addition, this result should be stored again in the database with the same value, and it should be possible to verify that this sequence of procedures has been performed well. At this time, it is inefficient to validate the entire data every time. Therefore, verify the integrity of the data by converting the results of each procedure into hash result values and comparing whether the hash result values are the same.

1.2. Public-key cryptography

Public-key cryptography is also called Asymmetric cryptography. In general, a user can generate a keypair between a public key and a private key using a keypair generation function, and it should not be possible to infer a private key with a public key. This generated key pair provides a scheme for 1) encrypting data with public keys and decrypting data with private keys, or 2) generating digital signatures of data with private keys and validating digital signatures with public keys.

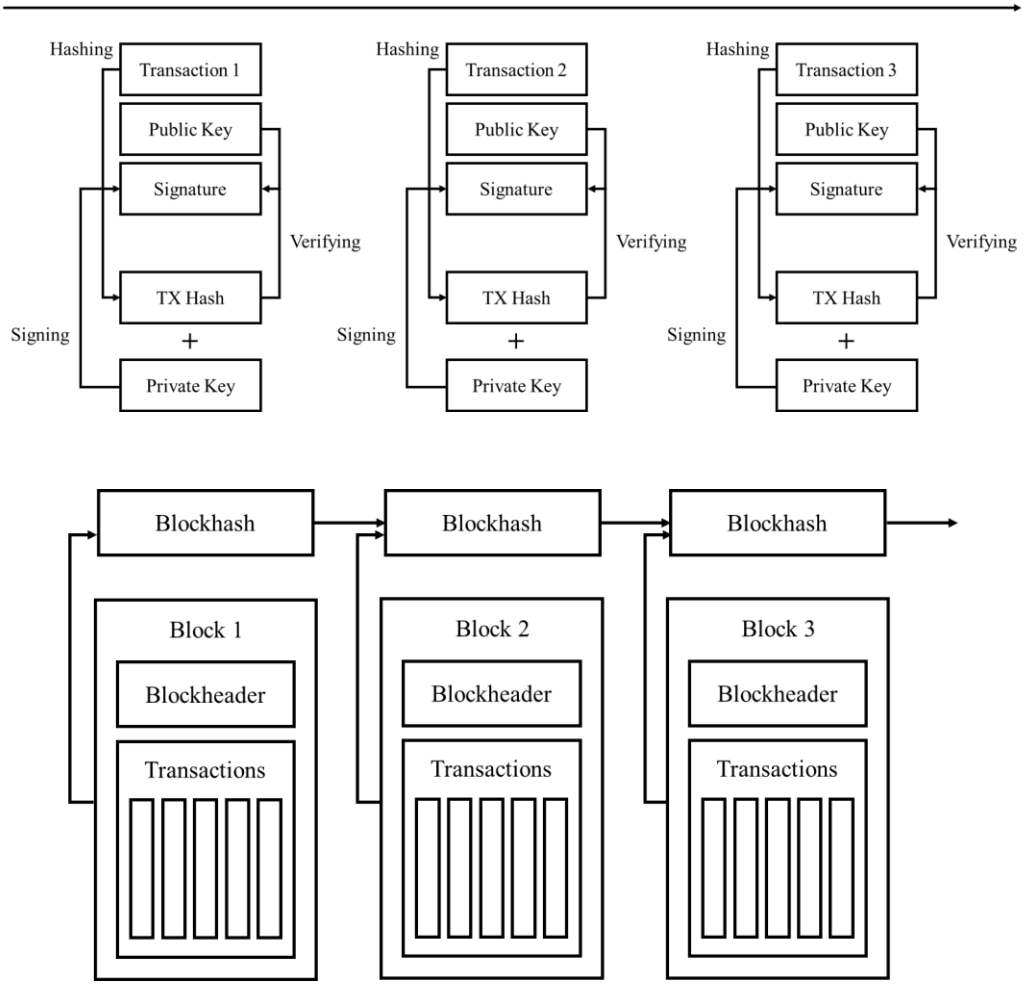


In blockchain, public-key cryptography is used to identify users using a digital signature scheme. Typically, the public key is hashed to generate an identifier called “Address”, and status data such as balances are attributed to the Address. By providing digital signatures, users of the blockchain network can register transactions, which are signs of intention to change status data attributed to each user, and using this series of procedures, concepts such as remittance, sales, and purchase can be implemented on the blockchain network.

1.3. Blockchain

A data storage system that combines data to bundle blocks and implements blocks to have connectivity with previous blocks using hash functions is called a blockchain. A type of connection in which multiple users share one blockchain data using public-key cryptography is called a blockchain network.

In a blockchain network, users can use the network by including transactions signed by individual keypairs in a block and synchronizing data. Since each block has connectivity with the previous block, if a specific user modifies the intermediate data of the blockchain, the integrity of the block data is destroyed and is not recognized as the correct blockchain data.



Each block has a block hash calculated from the previous block hash and block data, and users of the blockchain

network can see that the data is synchronized well by checking whether the block hash value is the same without comparing all data. Manipulating intermediate block data requires finding values that can cause hash collisions while passing all validations, which is almost impossible with current computing power.

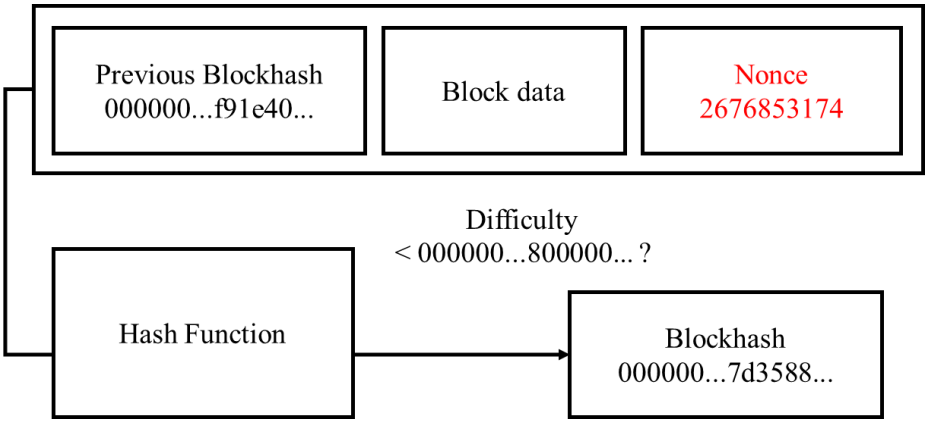
In general, attacks that manipulate block data are more economical to generate and overwrite all block data from the middle to the present. Therefore, blockchain networks have algorithms that prevent block data from being generated in a short time, and there is Proof of Work as a representative method.

2. Proof of Work

If a certain amount of computing power must be used to generate block data, an attacker who wants to modulate the data must inject as much computing power into block data generation at a sunk cost to modulate the blockchain data.

Earlier, it was revealed that it was difficult to find the original input value with the hash result value of the hash function. If there is a rule to find the original input with a specific hash result as a condition for generating a block and there is no valid attack method of the hash function, network participants have no choice but to find the original input through random substitution.

The act of network participants using computing power to find the original input of hash results for new block data is called mining, and the more computers participate in mining, the less likely the data in the blockchain network is to be falsified. It is common to provide sufficient incentives because there is no reason for participants to voluntarily consume computing power.



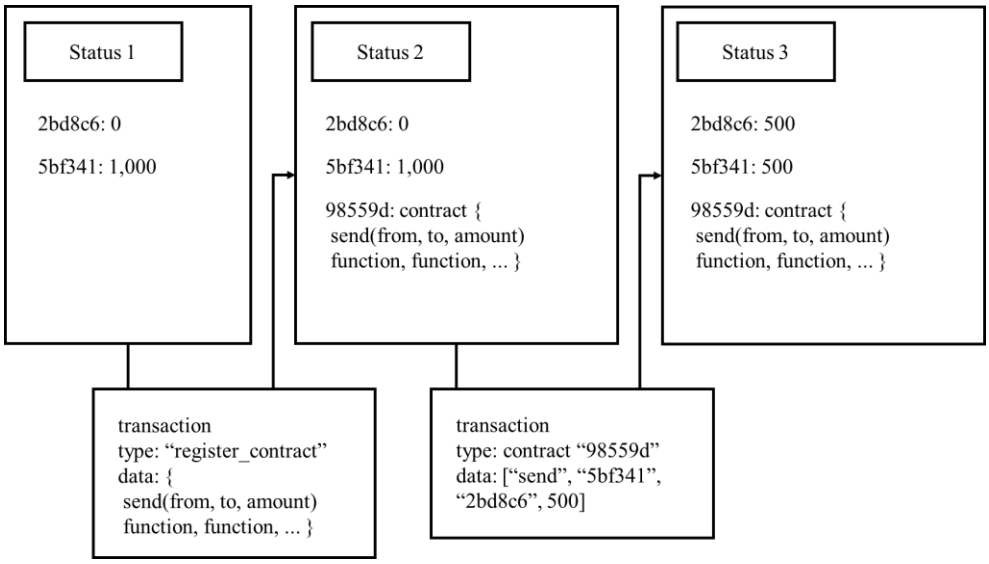
Since it is almost impossible to find the exact original input with a specific hash result value, finding the original input with a hash result value smaller than a certain value allows the block data to be connected. The probability of finding a hash that starts with 0 in the hexadecimal hash is 1/16, the probability of finding a hash that starts with 00 is 1/256, the probability of finding a hash that starts with 000 to 008 is 1/512.

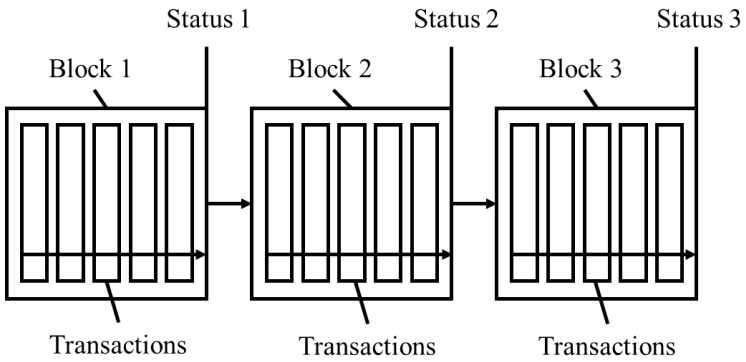
On average, when designing a block to be generated one in 10 minutes, the difficulty level can be set to 1/2 the probability of finding the hash in the next block generation if the previous block was generated in 5 minutes, and the difficulty level can be set to 2 times the probability of finding the hash in the next block generation if the previous block was generated in 20 minutes. Frequent changes to the difficulty level hinder the stability of the network, so it is generally designed to change the difficulty periodically over a period of 1 to 14 days.

3. Smart Contract

Satoshi Nakamoto proved that the combination of blockchain and proof of work can fully implement an electronic cash system without a reliable third-party. Vitalik Buterin, in an attempt to apply blockchain technology to other fields than currency, presented the concept of "Smart Contract" that contains executable code in block data.

Network users can register or modify new code by calling code that manages code, and users can use new code distributed by themselves or others by calling registered code. To this end, the type of data stored in the blockchain is divided into Transaction and Status, and by sequentially calculating Transaction, the status of the blockchain network can be finally obtained.





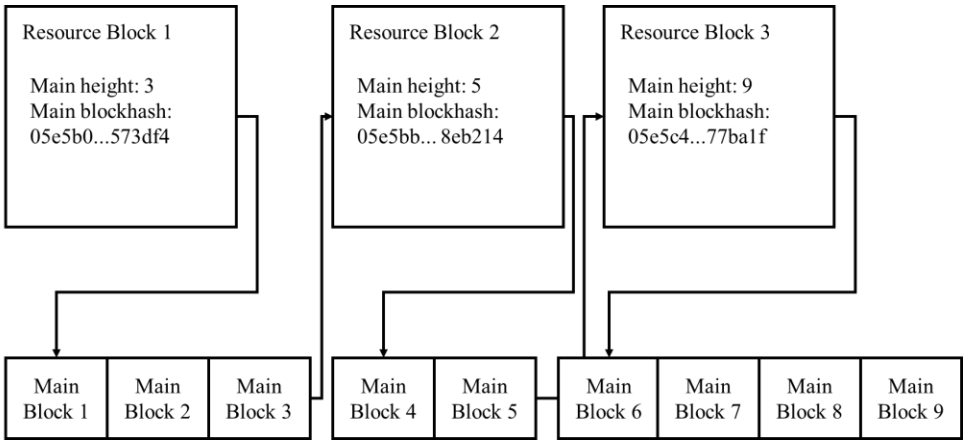
The system provides a contract, Virtual Machine, which manages code such as Genesis and Register, Modify, Delete. Other common service concepts are designed as contracts and incorporated into block data. Virtual Machine specifies areas of Status data that code registrants and users can access and manipulate to prevent incorrect code execution and limit code execution time to prevent code execution indefinitely.

4. Validators

The Proof of Work based blockchain system has a disadvantage in that it cannot guarantee transaction registration completeness in a short time. Blockchain researchers tried to solve this disadvantage by designating Validator through a certain method. However, there is a problem that the method that relies entirely on the Validator alone has no sunk cost for block data generation, so that all block data from the middle to the present can be newly generated and overwritten through double signatures.

4.1. Dual Chain

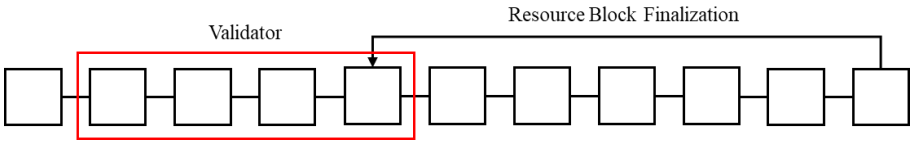
Since an algorithm that relies solely on Validator is dangerous, it should be used in combination with the algorithm of proof of work, and this system is called Dual Chain. Dual Chain consists of a Resource Chain that selects a Validator through proof of work and a Main Chain that processes general transactions as a verification algorithm by Validator.



Users who succeed in mining get some incentives and are elected as the next Validator, and need to fully process the general transaction and create the Main Block to obtain greater incentives. Network participants synchronize the Resource Chain based on the longest Resource Block, and synchronize the Main Chain based on the block hash of the Main Chain recorded in the Resource Chain. Block data in the Main Chain, not recorded in the Resource Chain, is considered provisional unless a separate branch is observed, thus ensuring fast transaction registration completeness unless malicious users are observed in real time.

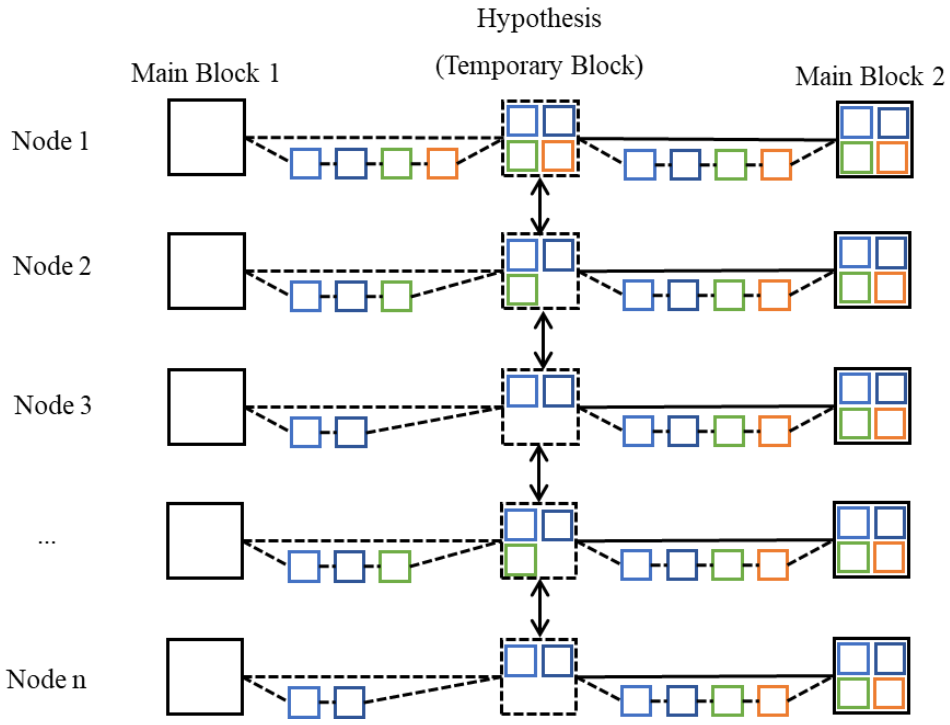
4.2. Hypothesis Acceptance Procedure

With the introduction of Dual Chain, it is possible to maintain decentralization and ensure fast processing of transactions. However, there is a risk that participants who succeeded in mining and were selected as Validators may temporarily stop or interfere with the network.



To prevent this situation, participants who succeed in mining are not excluded from the Validator immediately when the next miner appears, but are required to perform the role of Validator for a certain period of time to maintain a number of Validators. In addition, a number of selected Validators use Hypothesis Acceptance Procedure, an algorithm that allows multiple Validators to participate in block creation at the same time, rather than the Round-robin. This procedure is as follows.

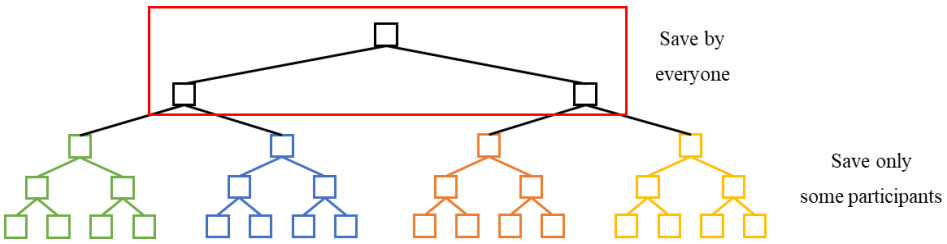
1) Validators generate a hypothesis with their own signature attached to the Blockheader of the block to be generated.
2) All Validators synchronize hypothesis data with each other.
3) If different hypothesis data exist, the best hypothesis is selected according to certain rules.
4) Validator, who presented not the best hypothesis, broadcasts it with a signature supporting the best hypothesis.
5) As a result of synchronization, if more than 66% of Validators support the same hypothesis, the block is determined.



The biggest advantage of the Hypothesis Acceptance Procedure method is that the location (IP) of the Validator is not specified because the block is determined only by the signature of the hypothesis. This helps blockchain networks to have strong resistance to DDoS attacks.

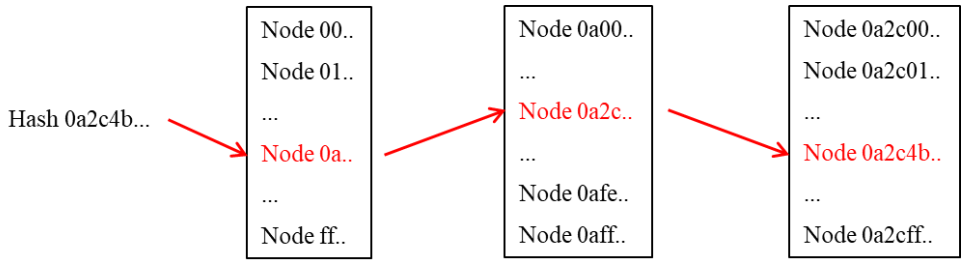
5. Partial Data Storage

When transactions are processed at high speed, the rate of increase in overall data capacity also increases rapidly. In traditional blockchains, it is not a big problem because hardware performance will increase by Moore's law, but in high-speed blockchain networks such as SASEUL, the capacity of the entire data exceeds the limits that a single computer can store.



Network participants separate blockchain data into data to be stored in common and data to be stored individually. After, participants save only data to be stored according to their Peer Address, and delete the remaining data. In general, the Peer Address is determined randomly, so it will be evenly distributed when the number of network participants increases. In addition, participants who want the blockchain network to remain intact will build multiple nodes to store the entire data.

As described above, blockchain data is divided into two types of transaction data and status data. Since the transaction data is included in the block data, it can be organized into a Merkle tree to classify the common and individual storage parts. Status data consists of a Merkle tree using a Hashtable in the form of a Key-Value to classify the common and the individual storage part.



If the data is divided into too many parts, the peer that holds the data cannot be found immediately at once. In this case, the peer that is storing data is found in a manner similar to the principle of finding an IP address.

6. Universal Computing

6.1. Proof of Observation

The act of finding Nonce in Proof of Work can be seen as proof that it consumed computing power. In SASEUL, not only miners but also network participants who retain a large amount of data have become important. Therefore, incentives should be provided to participants who are providing sufficient storage capacity.

When Validator processes a transaction, it must query the data of network participants with partial data if they do not have full ledger. When the Validator validates the state and the block data through neighboring nodes in the process of generating the Main Chain, the party being inquired receives the signature of the Validator, which is called a Receipt. Although it is not possible to verify whether the Receipt gave the data validly, it is evidence that the Validator inquired the data, and the peer holding the correct data will be able to collect a large number of Receipts.

Receipt is converted to an incentive as soon as it is recorded in the Resource Chain, and after creating a block of the Resource Chain, the Validator will mine the past Receipt while sharing some of the incentives converted from it. Increasing storage capacity is a more efficient means of mining because storage is generally much cheaper than CPU

or GPU.

6.2. Multi-chaining

Interworking while maintaining the data integrity of different programs or blockchain networks is called Multi-chain, which is not difficult. The following is an example of a procedure for implementing Swap between different blockchains without a bridge server.

1) (Chain A) X deposits 50A. Deposited amounts can be converted as Approve transactions when Y enters the original value of a specific hash. However, Y only knows the result value of the hash, not the original value. This status can be canceled after 10 minutes.
2) (Chain B) Y deposits 100B. Deposited amount can be converted as Approve transaction when X enters the original value of the hash generated in step 1). Since Y knows the hash result value in Chain A, it is possible to set a contract in Chain B that is based on the original value of Chain A, although it is a different chain. This status can be canceled after 5 minutes.
3) (Chain B) X enters the original value of the hash and completes the transaction of step 2).
4) (Chain A) Since the transaction ot step 2) has been completed, Y can know the original value of the hash generated in step 1).

By combining the Receipt and Multi-chaining techniques, it is possible to create not only a query of data but also a proof of the execution of the program. Also it is possible to contain a separate executable binary code in the block, not a contract code. In other words, depending on the design, all network participants can be a key element of distributed computing, which is called Universal Computing.

7. Random Number

In a blockchain network, a conventional random number system cannot be used because all participants must calculate the same result. The SASEUL-based blockchain network provides a "Random Number" variable within the contract, which is determined by the block hash at the moment the block is created. In SASEUL-based networks, the block hash value of the main chain is almost unpredictable, thus providing a complete random number within the maximum computational limit of the block hash.

8. Conclusion

In this article, I present 1) how blockchain networks process blocks at high speed, 2) how to distribute and process infinitely growing blockchain data and 3) how incentives generated while maintaining blockchain networks can be distributed through indirect mining. However, the methods proposed by SASEUL are only a starting point for Universal Computer implementation, leaving room for further study in the future.

Reference

[1] Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System” https://bitcoin.org/bitcoin.pdf

[2] Vitalik Buterin, “A Next-Generation Smart Contract and Decentralized Application Platform”

https://ethereum.org/669c9e2e2027310b6b3cdce6e1c52962/Ethereum_Whitepaper_-_Buterin_2014.pdf

[3] Jae Kwon, “Tendermint: Consensus without Mining” https://tendermint.com/static/docs/tendermint.pdf