www.gsr.io
gsr@gsr.io

How Bitcoin Works

Highlights

First conceived as a peer-to-peer electronic cash system, Bitcoin and the technology it's built upon is in process of changing humanity as we know it. In our first long-form version of our Chart of the Week series, we walk through a semi-technical, yet accessible explanation of how Bitcoin works.

- ***Purpose:*** We are often asked by those new to crypto how Bitcoin works. Moreover, we strongly believe a basic knowledge of the technology underpinning Bitcoin provides an invaluable foundation for understanding cryptocurrencies more broadly. As such, we attempt to explain how Bitcoin works in a semi-technical, yet accessible manner, focusing on what actually happens when a user initiates a Bitcoin transaction. Along the way, we cover various topics including hashing functions, public and private keys, digital signature algorithms, transaction validation, nodes and miners, unspent transaction outputs, proof of work consensus, the mining puzzle, the longest chain rule, transaction speed and finality, Merkle roots, and blockchain structure, among others.
- ***More than Just Electronic Cash:*** Satoshi Nakamoto, the anonymous developer of Bitcoin, first released the Bitcoin whitepaper in 2008, describing his or her vision for a "peer-to-peer electronic cash system." While Bitcoin has evolved into a store of value given technical limitations and a fixed supply, the technology behind Bitcoin has been extended to payments, identity, ownership, decentralized computation, etc., and brings along many positive attributes such as decentralization, censorship resistance, immutability, permissionlessness, and pseudonymity. Indeed, while fintech innovates on the front-end, it still operates on existing, often antiquated rails and simply exchanges one intermediary for another. Blockchain technology, by contrast, innovates on the back-end, reimagining the rails themselves and removing the intermediaries altogether.
- ***A Bird's Eye View:*** At a high level, the Bitcoin blockchain can be thought of as a decentralized database or distributed ledger comprised of a network of computers, often called nodes and miners. To send a bitcoin, an individual uses an application that takes the transaction information, the recipient's public key, and their own private and public keys to generate, digitally sign, and broadcast the transaction to the network. Once the network receives the transaction, miners process it by running validation tasks like checking that the signature is valid and organizing transactions into blocks. To ensure all nodes/miners have the same valid copy of the distributed

ledger, a consensus mechanism is used to determine which miner gets to post their block to the blockchain. Bitcoin uses a proof of work consensus mechanism, where miners work to be the first to solve a puzzle, with the winner earning the right to post the block and receive the block reward and transaction fees. The distributed ledger then stores the data in a chained format using hash pointers to make the blockchain easily searchable and tamper-evident. We expand on each of these steps in the body of the report.

- ***The Final Word:*** The Bitcoin blockchain is a canonical list of all transactions ever recorded by the network, but at its very core, is simply a linked list - a linked list whose technology may one day form the basis to bank the unbanked, democratize value exchange, remove rent-extracting intermediaries, establish new ownership and governance paradigms, and reshape the world as we know it.

Author: Brian Rudick, Senior Strategist

Table of Contents

Details

More than Just Electronic Cash

Satoshi Nakamoto, the anonymous developer of Bitcoin, first released the bitcoin whitepaper in 2008, describing his or her vision for a "peer-to-peer electronic cash system." While bitcoin has evolved into a store of value given technical limitations and a fixed supply, both of which are discussed below, the technology behind bitcoin has been extended to myriad use cases such as payments, identity, ownership, decentralized computation, etc. Moreover, its construction brings about many positive properties such as decentralization, censorship resistance, immutability, permissionlessness, and pseudonymity. Unlike fintech, which innovates on the front-end, operates on antiquated financial rails, and simply exchanges one intermediary for another, blockchain technology, by contrast, innovates on the back-end, reimagines the rails themselves, and removes the intermediaries altogether.

Exhibit 1: Bitcoin Properties & Blockchain Use Cases

Bitcoin Properties	
Decentralized	Permissionless
Trustless	Open Source
Censorship Resistant	Pseudonymous
Immutable	Scarce

Blockchain Use Cases	
Store of Value	Ownership
Payments	Governance
Record of Account	Decentralized Computation
Identity	Many More

Source: GSR

In the sections that follow we walk through how Bitcoin's SHA-256 hashing function creates a digital fingerprint for a piece of data. We then review symmetric and asymmetric cryptographic methods before covering public and private keys and Bitcoin's Elliptic Curve Digital Signature Algorithm, both of which allow only the holder of the private key to sign a transaction and for the network to check the validity of the signature. Next, we look at how that transaction is broadcast to the network of nodes, and how miners arrange the transactions into blocks to win the right to add the block to the blockchain and receive the mining reward. And we examine how this proof of work consensus mechanism establishes trust between unknown participants without the use of an intermediary or central counterparty, creating a shared truth that is the blockchain. Finally, we look at the structure of the blockchain, with each block linked to the previous and subsequent block, that when combined with Bitcoin's decentralized network, make Bitcoin the most secure database to ever exist.

Exhibit 2: Anatomy of a Bitcoin Transaction

Party	Action
User	An individual uses an app such as a digital wallet to take transaction info, the recipient's public key, and their own private key to generate, digitally sign, and broadcast the transaction to the network
Nodes	Nodes receive, validate, and relay the transaction to other nodes on the network via a gossip protocol
Miners	Miners organize transactions into blocks, work to solve the mining puzzle, and once done, broadcast the winning block/puzzle solution to the network before its verified and added to the blockchain
Nodes	A copy of the blockchain is maintained and continuously updated on thousands of computers around the world. A transaction will be considered final once several transactions have occurred on top of it

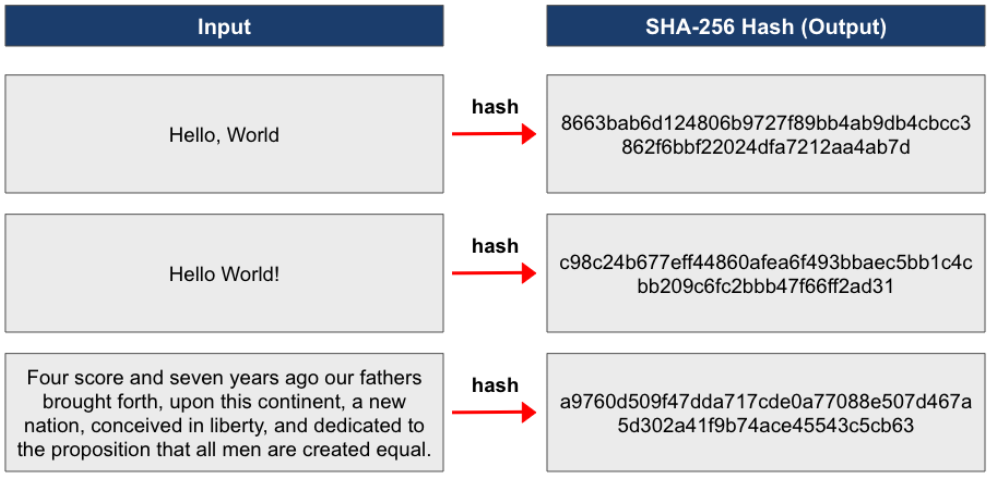
Source: GSR

Hashing Functions: Creating Data's Digital Fingerprint

Before moving further, we delve into cryptographic hashing functions, as they are used extensively in digital signatures, during the consensus process, and in the blockchain's construction. A hashing algorithm is a function that converts an arbitrary amount of data into a numeric string of fixed length, such as 256 ones and zeros or 64 hexadecimal characters, and can be thought of as that data's digital fingerprint. To create this unique identifier for the input data, the hash function splits the data into pieces and runs many rounds of local operations on them like AND, OR, and XOR, losing information as it goes. Hashing functions should be one-way (the only way to know the input from a given output is to try all possible inputs), deterministic (returns the same output for a given input), easy to compute (but not so easy that one can quickly cycle through all potential inputs to solve), and produce few collisions (two different inputs should not produce the same output).

Hashes have several benefits, such as improving efficiency and allowing for data verification without revealing the contents of the data. For example, rather than store passwords in a database that could potentially be hacked, a website can store hashed passwords. Then, when a user enters his or her password upon log-in, the website can simply take a hash of the entered password and compare it to its database of hashed passwords, materially enhancing security by not storing the passwords themselves (most websites modify this by adding a unique, user-specific random number to a user's password prior to hashing in what's called a salted hash. That way, if a hacker does get ahold of hashed passwords, the hacker can't simply use a dictionary of hashes of common passwords to figure out some of the simpler passwords). Bitcoin uses a specific hashing algorithm called SHA-256, which can be explored in this online SHA-256 hash calculator. We show several inputs and outputs in the exhibit below. Notice the high avalanche effect, where making one small change to the input data completely and unpredictably changes the resulting hash.

Exhibit 3: The SHA-256 Hashing Function



Source: GSR

Digital Signatures

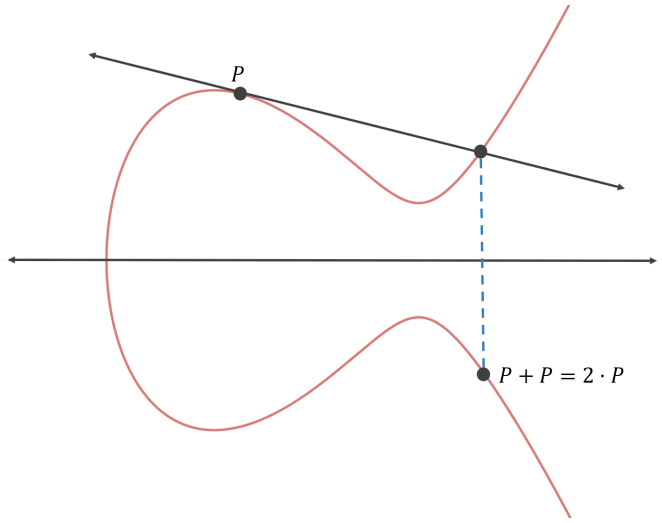
When designing a system of electronic money, it would appear at first glance that transactions could easily be faked, as a user could simply submit an invalid transaction or even copy a valid one. However, Bitcoin uses digital signature technology to, much like in the real world with physical signatures, ensure that only one specific person has the ability to sign a transaction. To do so, bitcoin uses public-key cryptography, also known as asymmetric cryptography.

Cryptology is the science concerned with data communication and storage in secure and usually secret form, encompassing both cryptography and cryptanalysis. One simple method of cryptography is symmetric cryptography, where two parties share the same secret key, which is used by the sender to encrypt the message and the recipient to decrypt the message. The Caesar Cipher is perhaps the simplest symmetric encryption technique and simply shifts the alphabet by a fixed amount, where, if the shared secret is three, A becomes C, B becomes D, and so on. The sender can encrypt the message using this shift, and the recipient can decrypt the message using this shift. While there are certainly more complex examples of symmetric cryptography than the Caesar Cipher and symmetric cryptography is generally less computationally intensive than more advanced forms of encryption, it requires a way to securely share the common secret key, which may be difficult or risky, and also requires a separate key for each pair of users in a network.

Asymmetric cryptography or public key cryptography, by contrast, utilizes mathematically linked public and private keys to eliminate the need to share a secret, and is more appropriate for large and expanding networks with frequent message sharing between different parties. Public keys, which are freely shareable, are created from private keys, which, as one-factor authentication mechanisms, should be kept strictly confidential. Importantly, the public and private keys are mathematically related, and while it is easy to calculate the public key from the private key, it is mathematically infeasible to go the other way. This property, combined with the ability to prove that one knows the private key without actually sharing it, enables the creation of digital secrets and signatures - allowing one to encrypt messages to a user's public key that can then only be decrypted by that user's private key, and allowing only the holder of a private key to sign a transaction.

Bitcoin uses a specific asymmetric cryptographic algorithm called an Elliptic Curve Digital Signature Algorithm, where a user selects a private key, usually at random, and runs elliptic curve operations on it to generate a mathematically linked public key that can't be used to infer the private key that created it. Elliptic curves follow the formula $y^2 = x^3 + ax + b$ (Bitcoin's is called secp256k1, which follows the curve $y^2 = x^3+7$), are symmetric about the x-axis, and any line drawn between two points will always intersect a third point. An elliptic curve cryptographic algorithm takes a starting point P, draws a line tangent to it, and takes the intersection point of that tangent line and the elliptic curve before flipping across the x-axis to generate a point 2*P (this set of operations is adding point P to itself). We can repeat this n number of times, cycling around the curve to end up at a point Q, as defined by Q=n*P. Q will seemingly have no relationship to our starting point P, and it is computationally infeasible for someone to know n (how many times you cycled around the curve) even when one knows the curve, Q and P. We can therefore use n as the private key and Q as the public key.

Exhibit 4: Elliptic Curve Digital Signature Algorithm



Source: GSR

In practice, both the private key and public key are abstracted from the user, and instead, a hash is taken of the public key to create a Bitcoin address (note that an address can also represent a script). The Bitcoin address serves as a unique identifier for the source or destination of a Bitcoin payment, is between 26 and 35 alphanumeric characters long, and is often converted to QR code format for ease of use. Bitcoin addresses provide an additional level of security because if there was ever a vulnerability in elliptic curves allowing someone to derive a private key from a public key, one's Bitcoin would still be safe since the public key isn't known until Bitcoin are spent (as the public key is required to validate the digital signature). Note that while we have transparency around addresses, in most cases we don't know who the address belongs to, which makes Bitcoin pseudonymous.

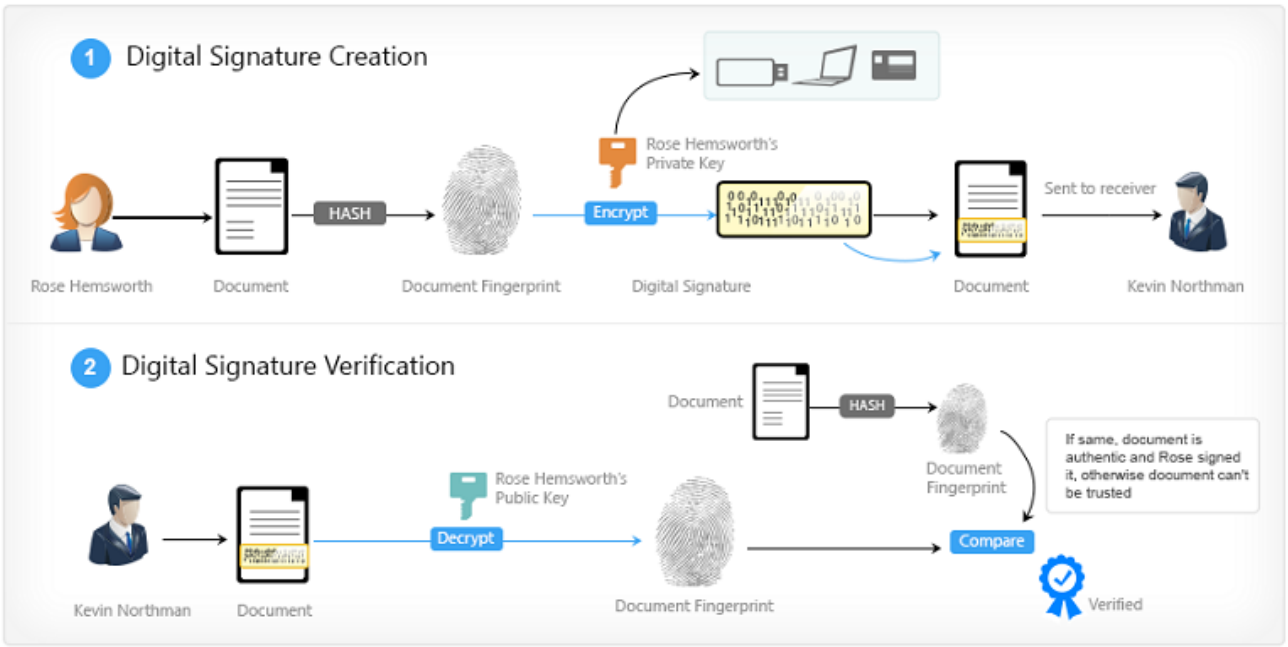
Exhibit 5: Bitcoin Address and Key Example

Source: GSR

Private keys are stored in a wallet, which don't actually store tokens but instead help facilitate any changes to the record of transactions stored on the blockchain. Wallets can take many forms, including paper wallets (keys written on a piece of paper), hardware wallets (keys stored on a thumb-drive), or online wallets (keys stored in an app or other software). Paper and hardware wallets are considered cold storage as they are not connected to the internet and are therefore more secure, though may not be as convenient and are susceptible to loss.

To create a digital signature, a user takes a hash of the transaction being signed. After the data is hashed, the user signs it using the ECDSA with transaction information, a random number called a nonce, and their private key as inputs to generate a digital signature consisting of two coordinates. This digital signature can then be cryptographically verified (ie. proven that it came only from the person holding the private key) using only the digital signature, the transaction, and the sender's public key. One does not need the private key to verify the transaction, and, since the digital signature depends on a nonce and the transaction itself, one's digital signature will be different for every transaction, preventing malicious actors from simply copying prior valid transactions.

Exhibit 6: Public Key Cryptography



Source: BitcoinClassroom.org, GSR

The Role of Nodes

Once the transaction is signed, a user will broadcast that transaction to a peer-to-peer network of computers called nodes that are responsible for validating transactions, sharing pending transactions and confirmed blocks with other nodes, enforcing the rules of the network, and maintaining a local copy of the blockchain on their machines. Upon hearing a newly broadcast transaction from a nearby user, a node will check its validity by making sure it followed a prescribed set of rules such as having a valid signature and making sure the sender had enough Bitcoin for the transaction. If the transaction checks out, the node who heard the transaction will then send it to the nodes closest to it in terms of latency, and the process continues until all the nodes on the network are aware of the pending transaction. Propagating transactions to the network through a gossip protocol such as this is more efficient than if transactions were broadcast to the entire network all at once.

Nodes are not compensated for their work, but running a node allows for certain privacy and security benefits for its operator, and more importantly, nodes play a vital role in network security and decentralization. Indeed, by keeping a full, essentially real-time copy of the blockchain on thousands of computers all around the world - there's even a node in space - the blockchain cannot be taken down, and, together with its consensus mechanism and chained structure discussed momentarily, bring about the security and decentralization attributes that are so important. Note that nodes may keep a full copy of the Bitcoin blockchain, known as a full node, or may only hold block headers, known as a light node. Light nodes help keep miners in check, are able to verify the existence of a transaction after receiving a limited amount of information from a full node, and increase decentralization, as their lower processing and storage requirements lower barriers to network participation.

Exhibit 7: Bitcoin Global Nodes Distribution

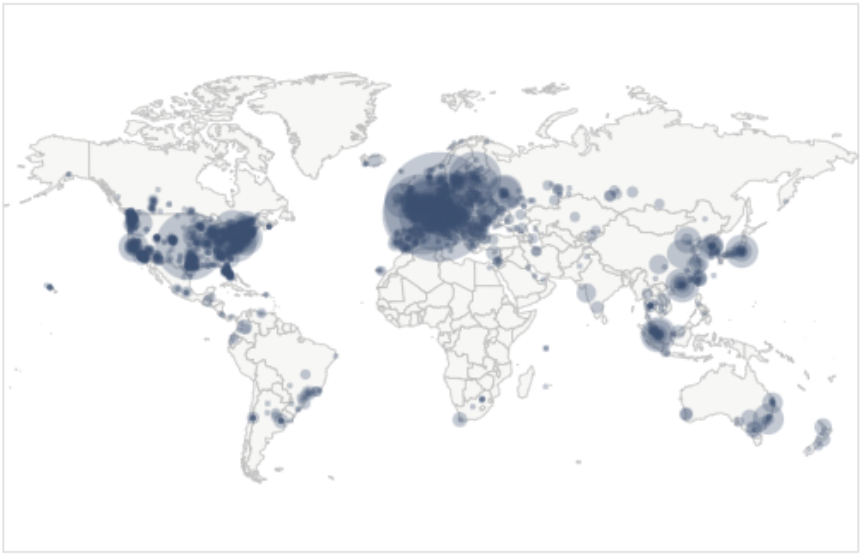
GLOBAL BITCOIN NODES DISTRIBUTION

Reachable nodes as of Thu Sep 23 12:49:20 2021 EDT.

11198 NODES | 24h | 90d | 1y

Top 10 countries with their respective number of reachable nodes are as follow.

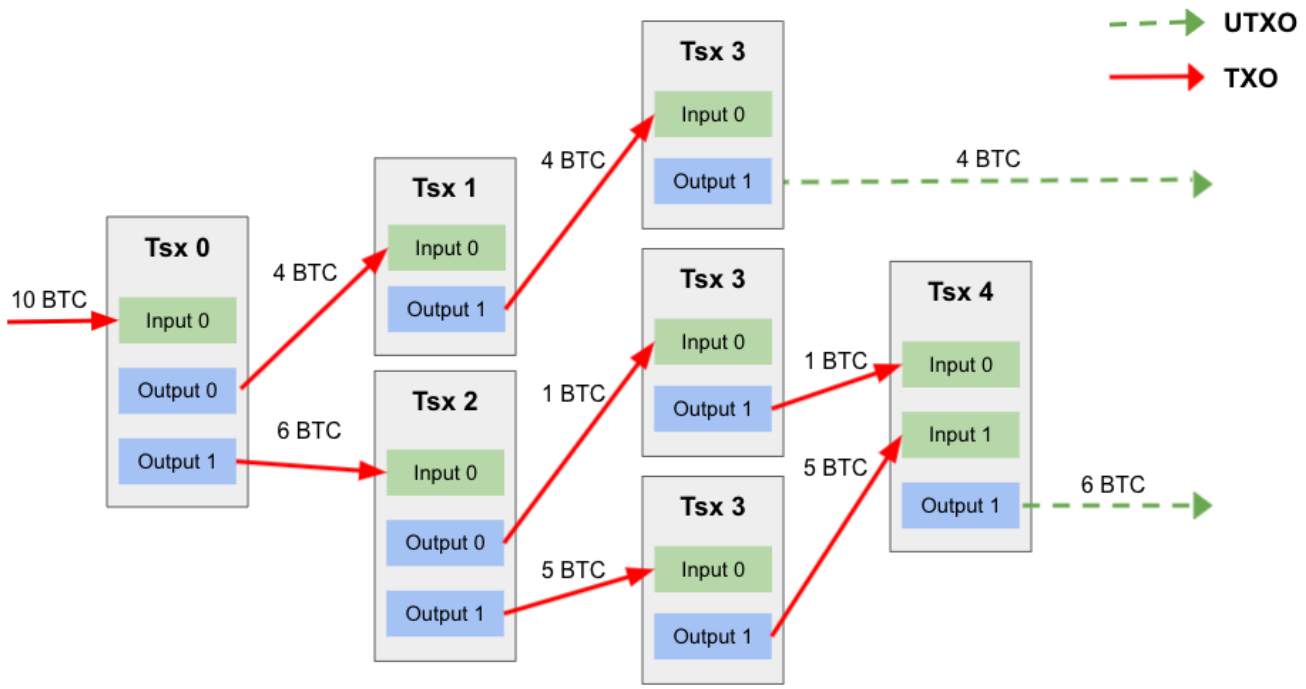
RANK	COUNTRY	NODES
1	n/a	3689 (32.94%)
2	United States	1872 (16.72%)
3	Germany	1807 (16.14%)
4	France	549 (4.90%)
5	Netherlands	384 (3.43%)
6	Canada	323 (2.88%)
7	United Kingdom	262 (2.34%)
8	Russian Federation	203 (1.81%)
9	Finland	186 (1.66%)
10	China	145 (1.29%)



Source: bitnodes.io, GSR

Lastly, note that unlike your bank account, Bitcoin is not an account-based system, but rather a collection of unspent transaction outputs, or UTXOs for short. UTXOs are transaction outputs that can be used as an input in a new transaction, require that new UTXOs equal the amount of old UTXOs that are being destroyed when the new UTXO is created, and represent a chain of ownership. While it sounds complicated, Bitcoin is simply keeping UTXOs that have come to a user separate, rather than summing into a total balance. As an example, if one has received a UTXO worth 0.1 BTC and a UTXO worth 0.5 BTC in historical transactions, and you wanted to send 0.2 BTC to a friend, you'd use the UTXO worth 0.5 BTC and send a 0.2 BTC UTXO to your friend and 0.3 BTC UTXO to yourself, which destroy the old 0.5 BTC UTXO (this simple example ignores transaction costs). UTXOs provide scalability, since multiple UTXOs can be processed in parallel, greater privacy when users create a new address for a transaction, and increased security and trust in the system as UTXOs can be traced all the way back to when Bitcoin was created as a block reward. Rules around how UTXOs can be spent are dictated by locking and unlocking scripts in the Bitcoin programming language Bitcoin Script. An unlocking script is part of every transaction input and contains the public address and digital signature discussed above, allowing the network to verify the spender of a UTXO is the holder of the private key.

Exhibit 8: Unspent Transaction Outputs (UTXO) Illustration



Source: GSR

Consensus in a Decentralized Network

Once a user has created a Bitcoin transaction and the transaction has been validated and propagated to the network, special nodes called miners collect and organize transactions into blocks. Much like in traditional finance, such as with ACH payments, this batching enhances efficiency. One issue that arises, however, is that the block each miner compiles is different from those of other miners, as a miner may include whichever transactions in a block that it would like up to the one megabyte block size limit (over the last year, blocks have generally averaged between 1,500 and 2,500 transactions per block). Moreover, miners will be aware of different transactions at different times due to network latency, also contributing to the differing blocks between miners. Further, since Bitcoin is an open network, there is the possibility that a bad actor will try to include a block with invalid transactions in it.

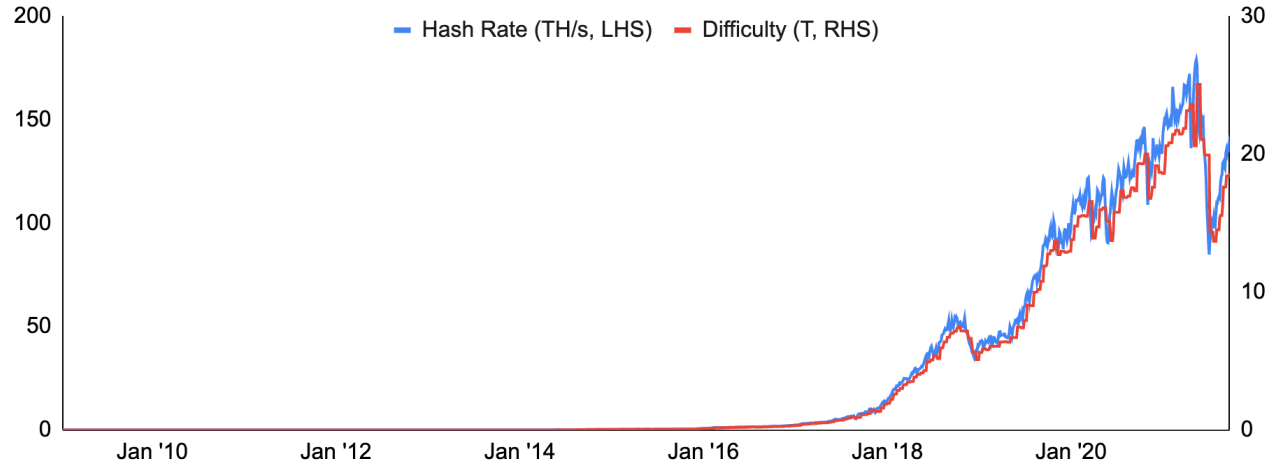
So how does a decentralized network composed of unknown parties agree on which transactions should go into a block and onto the blockchain, especially in the potential presence of bad actors? This is the question famously posed by the Byzantine Generals Problem, which is a mathematical representation of a battle with similar parameters. Specifically, there are several generals surrounding a city. If the generals all attack at the same time or all retreat at the same time, they will win. But, if the attack or retreat is not coordinated, they will lose. In addition, communication between the generals is poor, as some of the generals or messengers may be traitors and deliver false messages. Bitcoin solves this problem with its consensus mechanism, ensuring that all participants have a single shared truth despite no single coordinating party and the presence of potential bad actors (A more formal definition of consensus is an agreement on the state of the system or the transition between the current and next state, in Bitcoin represented by a set of valid transactions).

One might think that one way to accomplish this decentralized consensus would be to randomly select a miner, post their block to the blockchain, and as long as there is a certain threshold of honest nodes relative to dishonest nodes, then invalid blocks will be recognized and the honest nodes will build upon the valid chain. However, since Bitcoin is an open system and a malicious actor may create an unlimited number of miner nodes so they can continually be selected to post their invalid blocks, we need to choose the winning miner in a way that makes it costly for the miner, such as by requiring resources to participate and win. This is exactly the system Bitcoin's proof of work consensus mechanism employs, with its mining puzzle requiring significant computational and energy resources to solve. By requiring miners to have skin in the game in the form of considerable resources, and by having objective rules the system follows and a mechanism whereby malfeasance is immediately discoverable, Bitcoin encourages miners to participate honestly and solves the Byzantine Generals Problem, ie. is what's known as Byzantine Fault Tolerant or BFT.

Miners and Block Production

To solve the mining puzzle, miners will take their proposed block, attach a nonce (random number) to it, and hash it, in an attempt to achieve a certain number of leading zeros in the hash, which is the Bitcoin mining puzzle. While a good mining machine can cycle through nonces quickly, the time stamp will vary and miners can also vary the transactions/transaction order when attempting to solve the puzzle. Importantly, the probability that a miner is the first to solve the puzzle multiple times in a row, which would allow a malicious actor to continue posting blocks containing invalid transactions, is very low, making the network extremely secure. As a greater number of leading zeros requires more hashes to be run in order to solve and therefore takes more time, the Bitcoin network changes the number of leading zeros required to solve the puzzle as the number of miners and network hash power changes such that it takes roughly ten minutes for the network to produce a new block (technically, miners work to produce a hash that is less than or equal to a target, which, the more leading zeros there are, the more likely that this is the case. Changing the target rather than the required number of leading zeros allows the network to change difficulty in smaller increments). This adjustment to mining difficulty occurs every 2,016 blocks or roughly every two weeks. Notice the decline in difficulty in June this year after China's crackdown on mining and the resulting decline in difficulty. Then, as miners relocated and came back online, and total network hash power increased, network difficulty rose accordingly.

Exhibit 9: Bitcoin Mining Difficulty & Total Network Hash Rate



Source: Blockchain.com, GSR

As miners are repeatedly hashing data, specialized Application-Specific Integrated Circuits (ASICs) optimized for mining are used instead of CPUs or GPUs, resulting in reduced energy consumption and increased computing capacity.

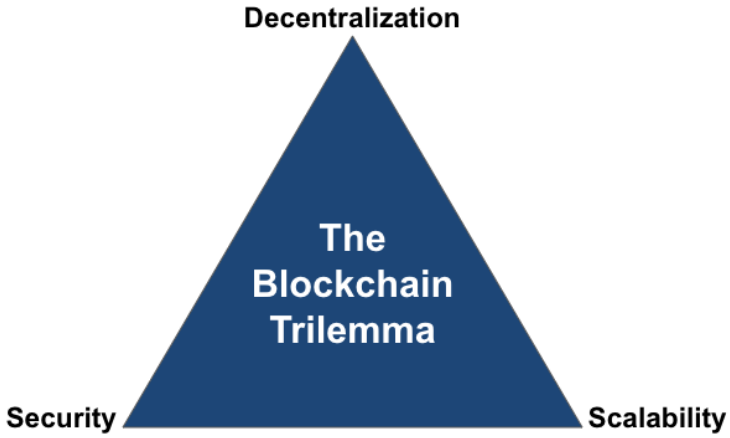
Exhibit 10: A Bitcoin Mining Facility

Source: GSR

Despite mining taking place on such specialized equipment, the limited size and frequency of blocks only allow the network to process roughly seven transactions per second. And while the block size could be increased to allow more transactions in each block or the time between blocks decreased, this would come at the expense of security and/or decentralization. For example, increasing the block size would lead to a larger blockchain, making it prohibitive for some nodes to keep a full copy, reducing decentralization. In fact, the difficulty/impossibility for a blockchain to achieve security, decentralization, and scalability is what Ethereum founder Vitalik Buterin coined as The Blockchain Trilemma, where he argues that blockchains can really only have two of the three attributes at once. Note that some layer two solutions claim to have solved The Blockchain Trilemma, as they are able to handle transactions off the main chain/layer one blockchain while relying on the security of the layer one blockchain itself. That said, there certainly are many nuances, and The Blockchain Trilemma does not appear to be solved with respect to all aspects, such as with data availability.

While blockchains secured by proof of work consensus mechanisms such as Bitcoin are extremely secure, they tend to have low throughput when compared to traditional financial companies or even to blockchains with different consensus mechanisms such as proof of stake that don't require such heavy computation. That said, tradeoffs exist, as proof of work generally has greater decentralization, less pre-mining, and lower inflation when compared to proof of stake consensus mechanisms. Bitcoin purposely chose to sacrifice scalability in return for decentralization and security, and its consensus mechanism allows miners to add transactions to the blockchain and fairly distribute the 21 million Bitcoin supply while adhering to these ideals.

Exhibit 11: The Blockchain Trilemma

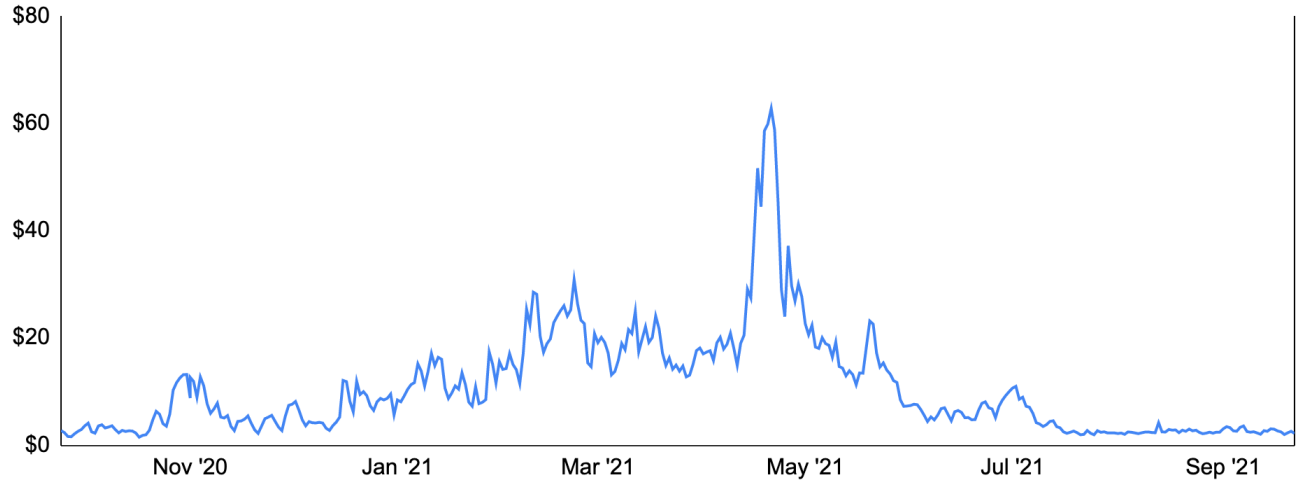


Source: GSR

One issue arising with limited throughput is that it leads to a high processing fee in the form of the "voluntary" miner tip, as users compete to get their transactions processed. This high cost per transaction is one reason why the Bitcoin network isn't frequently used for payments, in addition to its low throughput / slow finality and price volatility.

Note that this applies only to on-chain transactions, or transactions that are processed by the Bitcoin network and included in a valid finalized block. Transactions may also be done off of a given blockchain network, simply called "off-chain", and are often batched together before submitting back to the main chain at a later time. Bitcoin's Lightning Network and Liquid Network employ off-chain transactions, as do brokerages and exchanges. For example, when trading on a centralized exchange such as Binance or Coinbase, a user will initiate an on-chain transaction to send his or her Bitcoin to the exchange. The user's Bitcoin is then held by the exchange and he or she does not have access to the private keys. The user may then trade their Bitcoin as much as desired, and the exchange will record these transactions in its own private ledger, rather than sending any transactions to the Bitcoin network. Once a user requests to withdraw his or her Bitcoin, the exchange will calculate the net amount of all trades before initiating a second on-chain transaction to send the balance to the user. Off-chain transactions allow for scalability and typically offer lower fees and quicker settlement times, but are less secure than on-chain transactions and may have additional drawbacks such as capital requirements, slow finality, and/or centralized custody.

Exhibit 12: Average Fee per Transaction (USD), Last 12 Months



Source: Blockchain.com, GSR

As network difficulty has grown over the years, it has become more and more difficult for some, particularly smaller miners to consistently produce blocks, or even produce blocks at all. In response, mining pools were formed, where miners pool resources to mine blocks together, sharing processing power over a network and splitting the block reward in proportion to the amount of work contributed. Specifically, miners contribute hash rate, which is a measure of the number of hashes performed per second. When a miner in a pool mines a block, they send the block reward to the pool coordinator, who takes a small fee and pays each pool member in proportion to its hash rate contribution. Joining a mining pool allows miners to receive a portion of the block rewards on a more consistent basis, rather than randomly every few years. As mining benefits from economies of scale, the presence of mining pools allows Bitcoin mining to remain decentralized rather than amass to the largest of miners.

Exhibit 13: Mining Pool Distribution, Last One Month

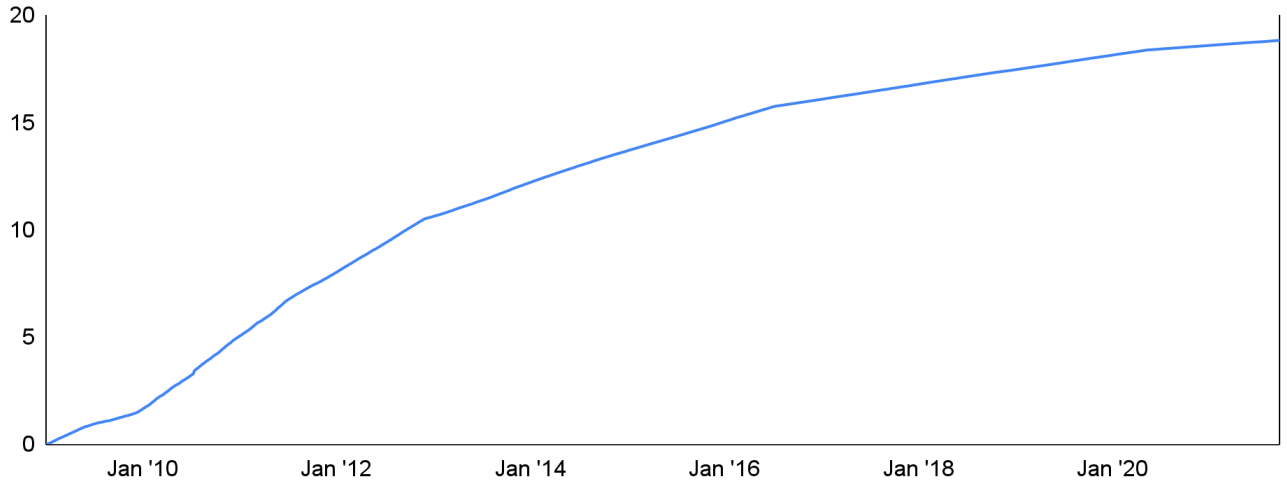
Rank	Network	Hash Rate Share	Blocks Mined	Empty Blocks Count	Empty Blocks %	Avg. Block Size (Bytes)	Avg. Tsx Fee per Block	Tx Fees % of Block Reward
1	AntPool	16.9%	753	1	0.1%	1,085,557	0.096	1.5%
2	F2Pool	14.7%	655	2	0.3%	1,089,117	0.092	1.5%
3	Poolin	13.1%	582	1	0.2%	1,096,924	0.090	1.5%
4	ViaBTC	12.6%	560	4	0.7%	1,069,031	0.092	1.5%
5	Binance Pool	10.0%	445	3	0.7%	1,108,582	0.097	1.6%
6	Foundry USA	8.8%	390	0	0.0%	1,104,425	0.087	1.4%
7	BTC.com	8.6%	384	0	0.0%	1,078,801	0.088	1.4%
8	SlushPool	4.9%	216	1	0.5%	1,081,053	0.086	1.4%
9	unknown	3.9%	173	0	0.0%	1,143,480	0.105	1.7%
10	Huobi.pool	1.8%	79	0	0.0%	1,050,916	0.093	1.5%

Source: BTC.com, GSR

Structure of the Bitcoin Blockchain

Once a miner solves the mining puzzle with a valid block, it will broadcast that block and the nonce that solved the puzzle to other miners, who then verify the solution. If a majority of miners reach consensus on the solution, the winning miner is allowed to add their block to the blockchain and receive the block reward and transaction fees. Block rewards are currently set at 6.25 BTC and are cut in half every 210,000 blocks or roughly every four years. Such halving ensures a fixed supply, as there will only ever be 21 million Bitcoin mined. And as block rewards go down, fees are likely to increase, though there are offsets that can sustain miner profitability and thus network participation, such as if the price of Bitcoin increases or if the price of electricity and mining equipment falls.

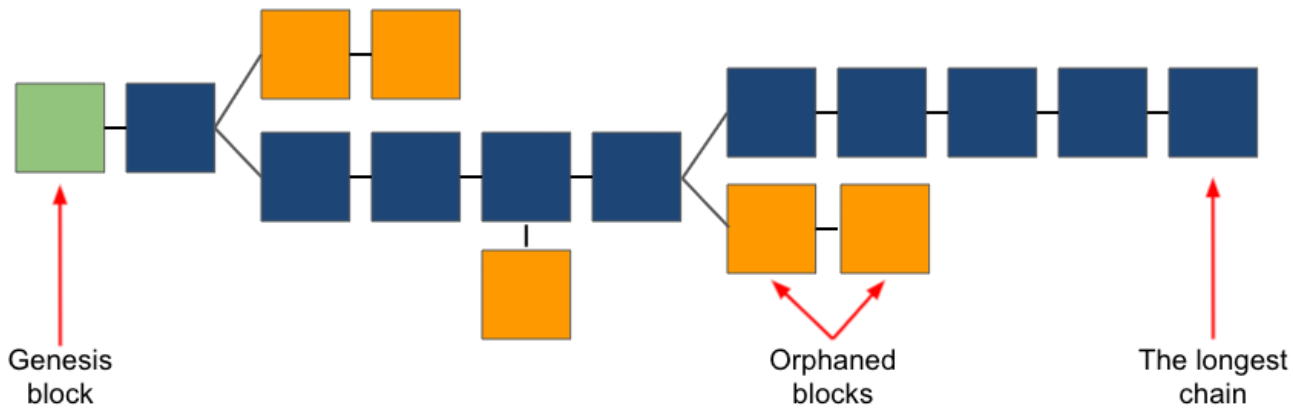
Exhibit 14: Bitcoin Supply, Millions



Source: Blockchain.com, GSR

One issue with mining is that every so often, two miners will solve the mining puzzle at roughly the same time. As these are both valid blocks, the blockchain will temporarily break into two chains in a process known as forking. Moreover, given network latency, some miners will hear about one winner and other miners will hear about another winner, at which point they both immediately move onto the next block attempting to build upon the block that it heard was the winner. Over time, the winner heard by the greatest number of miners will end up being the longest chain, which can be approximated by the number of blocks but technically is the chain with the most cumulative “chainwork”, as chains can span multiple difficulty periods. Regardless, shorter chains will eventually be abandoned, and the network will accept the longest chain as the main chain.

Exhibit 15: The Longest Chain Rule



Source: GSR

Such a construct ensures that the chain with the most work put into it is adopted and makes it very hard for a malicious actor to maintain a chain comprised of bogus transactions and blocks for very long. Valid transactions that are part of an orphaned block on an abandoned chain are simply added back to the pending transactions pool known as the mempool and are again available to be added to a new block by miners. Given this risk, many crypto providers require a certain number of confirmations to consider a transaction/block final. Confirmations are the number of blocks built upon the block containing the transaction in question. For Bitcoin, a transaction with six confirmations, which includes the block containing the original transaction, equates to a likelihood of 99.9% that block and associated transactions won't be orphaned and are part of what ultimately is the main chain.

Exhibit 16: Cryptocurrency Finality per Kraken's Confirmations Requirements

Asset	Confirmations Required	Estimated Time (minutes)
Bitcoin	4	40
Ethereum	20	5
Algorand	10	0.8
Cardano	15	10

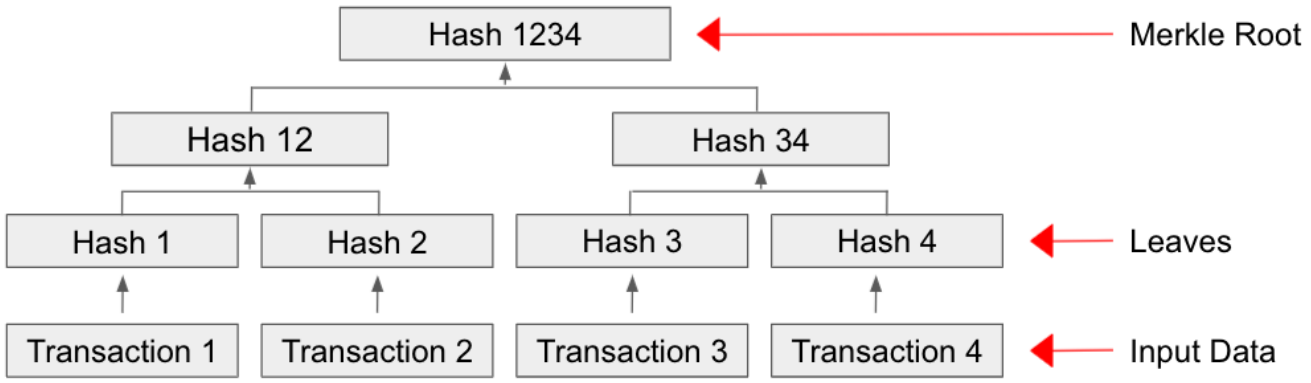
Asset	Confirmations Required	Estimated Time (minutes)
Flow	30	1
Tether	20	5
Polygon	20	5
Ethereum Classic	40,000	9,360

Source: Kraken, GSR

We have now covered how transactions are created, validated, propagated, arranged into blocks, and added to the chain. Just as important, however, is the specific structure of the blocks. Blocks, which are limited in size to one megabyte of data, contain two principal components, a block header and transaction data. The block header contains several pieces of information or metadata about the block, including the nonce that solved the mining puzzle, the hash of the previous block header, and a Merkle tree root.

Merkle trees are an elegant way to create a fingerprint of all the transactions in a specific block, and are best explained using an example. To generate the Merkle root for a block with four transactions, 1, 2, 3, and 4, miners simply take a hash of each transaction (technically a double hash), concatenate the hash of the first two transactions (hash 1 and hash 2) before taking a hash of it, which we'll call hash 12. Next, it'll concatenate the hashes of transactions 3 and 4 (hash 3 and hash 4) and take a hash of that, hash 34. Next, it will concatenate hash 12 and hash 34 before taking a hash of that, which is the Merkle root. Including the Merkle root in the block header has several advantages. First, if any part of any transaction is altered in the block, the Merkle root will change, making the block tamper-evident. In addition, Merkle trees allow light nodes, which only have the block headers, to validate that a block contains a certain transaction without needing to maintain a copy of the entire blockchain by receiving just the block hash and Merkle path associated with the transaction from a full node for the transaction in question.

Exhibit 17: A Merkle Tree



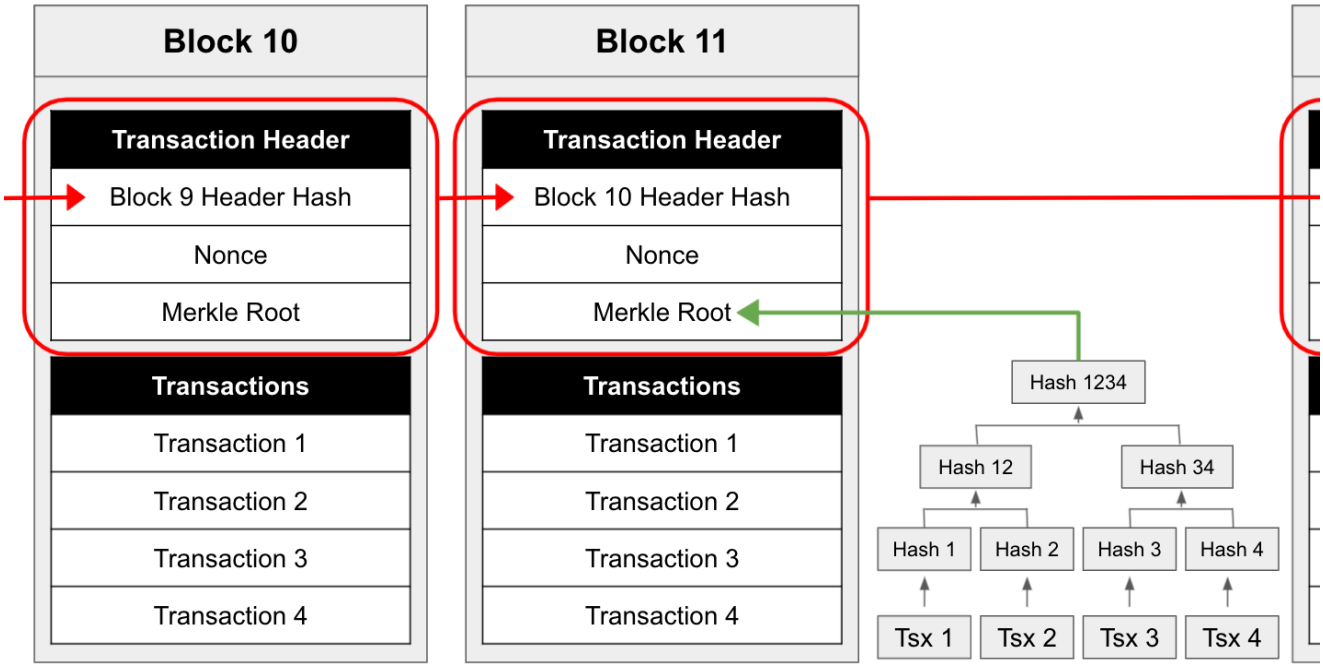
Source: GSR

As mentioned, the block header also includes a hash of the previous block header, which time-orders the blocks, provides for improved searchability, and makes them tamper-evident. This is because changing any historical transaction will change the Merkle root for that block, which, as it's included in the block header, will change the hash of that block header, which then won't match what was recorded for the prior block header hash in the subsequent block, breaking the chain.

Moreover, even if a bad actor changes a historical transaction, calculates a new Merkle root, and then

re-calculates the hash of the block header and inserts that block header into the next block, and continues in this fashion all the way through the current block, it can be easily observed that even though this individual blockchain foots with itself, it doesn't match the copy of the blockchain on all the other computers, and this can be observed simply by comparing the most recent block.

Exhibit 18: Bitcoin Block Structure



Source: GSR

We can now use a block explorer to search the Bitcoin blockchain for data about transactions, blocks addresses, and more. Below we show information for Block 701935 from Blockchain.com's Bitcoin block explorer, copied into a table for readability on one page. The block is split into two sections, the block header and block transactions. Inside the block header, we can see that the block was mined on September 23rd, 2021 by an unknown miner (though clicking on the link gives the address) and contains 1,129 transactions. We can see the hash of the block as well as the Merkle root of the transactions in the block. In addition, we can see the nonce that solved the mining puzzle (772,613,006), as well as the difficulty associated with the puzzle.

In the transaction section, we show two of the 1,129 transactions. The first is the block reward, or coinbase transaction, where Bitcoin was created to reward the miner for generating the block. We can also see a second transaction, the hash of which starts with 75a5b. Here, the address beginning bc1qw uses a UTXO of 0.12161 BTC to send 0.005 BTC as the transaction fee, 0.04269 BTC to the address starting with 1JdL, and, following the UTXO principles described previously, the remaining 0.0784 BTC back to itself.

Exhibit 19: Reading a Block Explorer: Block 701935

Block 701935		
Block Header		
Hash	000000000000000000e7e221139c35c9d6cc31959b2c174b8a83d04e9037606	
Previous Hash	000000000000000000bf6cc51327c245259db28642772b369d9b762e220cee4	
Confirmations	550	
Timestamp	2021-09-23 23:34	
Height	701935	
Miner	Unknown	
# of Transactions	1,129	
Difficulty	18,997,641,161,758.90	
Merkle root	b594703b70771d2b9c45ca49da1842225255c0b6d0dbb81b83da4cec69d3089a	
Version	0x20c00004	
Bits	386,846,955	
Weight	2,107,787 WU	
Size	765,722 bytes	
Nonce	772,613,006	
Transaction Volume	4649.53407463 BTC	
Block Reward	6.25000000 BTC	
Fee Reward	0.05035101 BTC	
Block Transactions		
Hash	17fd35509899ba64ef16e8ff07c725f9f2d9fa17557a46a487424e58651ad8e0	
Amount	6.30035101 BTC	
Fee	0.00000000 BTC	
Time	2021-09-23 23:34	
From	Coinbase Transaction	
To	19dENFt4wVwos6xtgwStA6n8bbA57WCS58	6.30035101 BTC
	OP_Return	0.000 BTC
Hash	75a5be6a963bad14d8bfbc8069f036a9a1ddf93572a4c4d279ab263dcf5b41b8	
Amount	0.12111004 BTC	
Fee	0.00050000 BTC	
Time	2021-09-23 23:28	
From	bc1qwqdg6squsna38e46795at95yu9atm8azzmyvckulcc7kytlcckxswvvzej	0.12161004 BTC
To	1JdLdLV7xqj3L2ZQXrhHe9UjDRscjTi2T6	0.04269902 BTC
To	bc1qwqdg6squsna38e46795at95yu9atm8azzmyvckulcc7kytlcckxswvvzej	0.07841102 BTC
1,127 more transactions		

Source: Blockchain.com, GSR

Conclusion

While it may seem complicated - and there's quite a lot we didn't cover - the process of sending a Bitcoin transaction is fairly straightforward. An individual uses an app such as a digital wallet to take transaction information, the recipient's public key, and their own private and public keys to generate, digitally sign, and broadcast the transaction to the network. Nodes receive, validate, and relay the transactions to other nodes on the network, before miners organize the transactions into blocks, working to solve the mining puzzle and add their block to the blockchain. Finally, a copy of the blockchain is maintained and continuously updated on thousands of computers around the world, and the transaction is considered final once several transactions have occurred on top of it.

At its very core, the Bitcoin blockchain is simply a linked list - a linked list offering invaluable attributes such as decentralization, trustlessness, permissionlessness, and immutability. One whose technology reimagines the financial rails themselves and removes intermediaries altogether. And one whose technology may one day form the basis to bank the unbanked, democratize value exchange, remove rent-extracting intermediaries, establish new ownership and governance paradigms, and reshape the world as we know it.

About GSR

GSR is a global leader in digital asset trading, market making, OTC derivatives, and investments. We operate in a culture of excellence and leverage our first-rate reputation, deep relationships and proprietary trading technology to move swiftly and capitalize on market opportunities.

GSR's experienced team brings together decades of institutional trading expertise, while our industry-leading proprietary technology stack anchors everything we do.

Our main service areas are market-making; proprietary and algorithmic trading; client execution; structured products; risk management solutions; and portfolio investments.

For more information or if we can help with anything, please see gsr.io or contact us at gsr@gsr.io.

Required Disclosures

This material is a product of the GSR Sales and Trading Department. It is not a product of a Research Department, not a research report, and not subject to all of the independence and disclosure standards applicable to research reports prepared pursuant to FINRA or CFTC research rules. This material is not independent of the Firm's proprietary interests, which may conflict with your interests. The Firm trades instruments discussed in this material for its own account. The author may have consulted with the Firm's traders and other personnel, who may have already traded based on the views expressed in this material, may trade contrary to the views expressed in this material, and may have positions in other instruments discussed herein. This material is intended only for institutional investors. Solely for purposes of the CFTC's rules and to the extent this material discusses derivatives, this material is a solicitation for entering into a derivatives transaction and should not be considered to be a derivatives research report.

This material is provided solely for informational purposes, is intended for your use only and does not constitute an offer or commitment, a solicitation of an offer or comment (except as noted for CFTC purposes), or any advice or recommendation, to enter into or conclude any transaction (whether on the indicative terms shown or otherwise), or to provide investment services in any state or country where such an offer or solicitation or provision would be illegal.

Information is based on sources considered to be reliable, but not guaranteed to be accurate or complete. Any opinions or estimates expressed herein reflect a judgment made as of the date of publication, and are subject to change without notice. Trading and investing in digital assets involves significant risks including price volatility and illiquidity and may not be suitable for all investors. GSR will not be liable whatsoever for any direct or consequential loss arising from the use of this Information. Copyright of this Information belongs to GSR. Neither this Information nor any copy thereof may be taken or rented or redistributed, directly or indirectly, without prior written permission of GSR. Not a solicitation to U.S. Entities or individuals for securities in any form. If you are such an entity, you must close this page.