A Gentle Introduction To Blockchain Technology

BRAVE NEWCOIN.
Digital Currency Insights

Table of Contents

A Gentle Introduction To Blockchain Technology

Authored By

Antony Lewis

Antony Lewis has a passion for virtual currencies such as bitcoin, and the underlying technologies behind them, including blockchain data structures and distributed consensus systems. Antony believes that these new ways of putting the technologies together will change the world of business, reminiscent of how the internet changed the distribution of information. Antony consults businesses, helping them understand the implications of blockchain technology.

antony@bitsonblocks.net

@antony_btc

Published By

BraveNewCoin

BNC. Digital Currency Insights

Adapted from

Bits On Blocks

The '**Gentle Introduction Reference Papers™** are the first in a series of accessible documents published by Brave New Coin for industry decision makers. Designed to demystify the inner workings of Bitcoin, Digital Currencies and the emerging Blockchain technology.

Series ONE covers:

"A Gentle Introduction To"

> Bitcoin

> Blockchain Technology

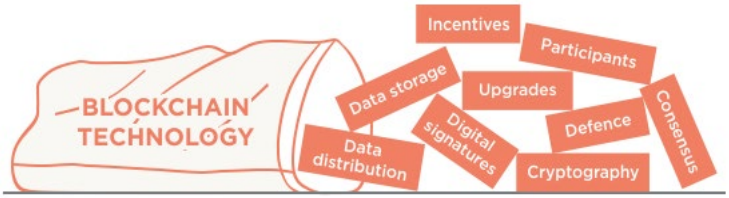
> Bitcoin Mining

> Digital Tokens

Free to Download and Share

Part 1

EXECUTIVE SUMMARY



Think of blockchain technology as a bag of Lego or bricks.

People use the term 'blockchain technology' to mean different things, and it can be confusing. Sometimes they are talking about The Bitcoin Blockchain, sometimes it's other virtual currencies, sometimes it's smart contracts. Most of the time though, they are talking about distributed ledgers, i.e. a list of transactions that is shared among a number of computers, rather than being stored on a central server.

The common themes seem to be a data store which:

- Usually contains **financial transactions**.
- Is replicated across **a number of systems** in almost **real-time**.
- Usually exists over a **peer-to-peer** network.
- Uses **cryptography** and **digital signatures** to prove identity, authenticity and enforce read/write access rights.
- Can be **written** by certain participants.
- Can be **read** by certain participants, a wider audience.
- Has mechanisms to make it **hard to change historical records**, or at least make it easy to detect when someone is trying to do so.

Think of "blockchain technology" as a collection of technologies, a bit like a bag of Lego. From the bag, you can take out different bricks and put them together in different ways to create different results.

Part 2

The Bitcoin Blockchain ecosystem

As a primer on bitcoin, it may help to review the original whitepaper by Satoshi Nakamoto titled Bitcoin: A Peer-to-Peer Electronic Cash System.

The Bitcoin Blockchain ecosystem is actually quite a complex system due to its dual aims: that anyone should be able to write to The Bitcoin Blockchain; and that there shouldn't be any centralised power or control. Relax these, and you don't need many of the convoluted mechanisms of Bitcoin.

Replicated databases. The Bitcoin Blockchain ecosystem acts like a network of replicated databases, each containing the same list of past bitcoin transactions. Important members of the network are called validators or nodes which pass around transaction data (payments) and block data (additions to the ledger). Each validator independently checks the payment and block data being passed around. There are rules in place to make the network operate as intended.

Bitcoin's complexity comes from its aims. The aim of bitcoin was to be decentralised, i.e. not have a point of control, and to be relatively anonymous. This has influenced how bitcoin has developed. Not all blockchain ecosystems need to have the same mechanisms, especially if participants can be identified and trusted to behave.

INTRODUCING BITCOIN'S BLOCKCHAIN

Here's how bitcoin approaches some of the decisions:

Category	Question	Bitcoin's approach	Other ways
Data storage	How should data be stored?	A blockchain	A database (could be replicated across multiple data centres)
Data distribution	How should new data be distributed?	Peer-to-peer	Client-server, hierarchical
Consensus mechanism	How should conflicts be resolved?	Longest chain rule	(Not needed in trusted networks) 'Trusted' or super-nodes
Upgrade mechanism	How do the rules get changed?	BIPs (for writing the rules) Vote by hashing power (for implementing the rules)	Centralised upgrades Contractual obligations
Participation criteria	Who can submit transactions?	Pseudonymous, open	Trusted, pre-vetted participants
Participation criteria	Who can read data?	Pseudonymous, open	Trusted, pre-vetted participants
Participation criteria	Who can validate transactions?	Pseudonymous, open	Trusted, pre-vetted participants
Participation criteria	Who can add blocks?	Pseudonymous, open	Trusted, pre-vetted participants
Defence mechanism	How to prevent bad behaviour?	Proof-of-work	(Not needed in trusted networks) Proof-of-stake, other 'proofs' or costs to add blocks
Incentivisation scheme	How to incentivise block-makers?	(only expensive in Bitcoin because of proof of work) Block reward, to be replaced by transaction fees	Contractual obligations 3rd party funding
Incentivisation scheme	How to incentivise blockchain data storage?	Not considered	Contractual obligations 3rd party funding
Incentivisation scheme	How to incentivise transaction validators?	Not considered	Contractual obligations 3rd party funding

Part 2

Public Vs Private Blockchains

There is a big difference in what technologies you need, depending on whether you allow ***anyone*** to write to your blockchain, or known, vetted participants. Bitcoin allows ***anyone*** to write to its ledger.

Public blockchains.

Ledgers can be 'public' in two senses:

1. Anyone, without permission granted by another authority, can write data

2. Anyone, without permission granted by another authority, can read data

Usually, when people talk about public blockchains, they mean anyone-can-write.

Private blockchains.

Conversely, a 'private' blockchain network is where the participants are known and trusted: for example, an industry group, or a group of companies owned by an umbrella company.

Many of the mechanisms aren't needed – or rather they are replaced with legal contracts.

This changes the technical decisions as to which bricks are used to build the solution.

Because bitcoin is designed as a 'anyone-can-write' blockchain, where participants aren't vetted and can add to the ledger without needing approval, it needs ways of arbitrating discrepancies (there is no 'boss' to decide), and defence mechanisms against attacks (anyone can misbehave with relative impunity, if there is a financial incentive to do so). These create cost and complexity to running this blockchain.

Part 3

DATA STORAGE: *What is a blockchain?*

A blockchain is just a file. A blockchain by itself is just a data structure. That is, how data is logically put together and stored. Other data structures are databases (rows, columns, tables), text files, comma separated values (csv), images, lists, and so on. You can think of a blockchain competing most closely with a database.

Blocks in a chain = pages in a book
For analogy, a book is a chain of pages. Each page in a book contains:

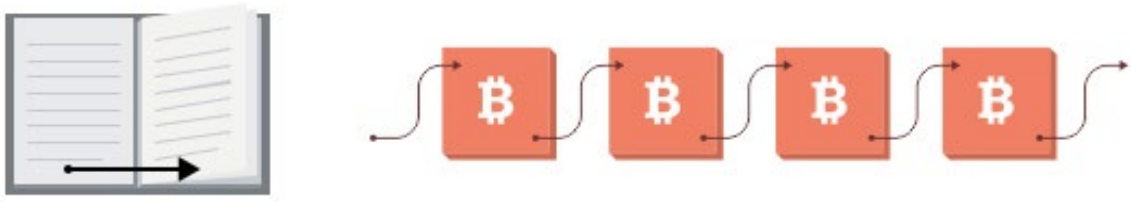
- **The text:** for example the story.

- **Information about itself:** at the top of the page there is usually the title of the book and sometimes the chapter number or title; at the bottom is usually the page number which tells you where you are in the book. This 'data about data' is called meta-data.

Similarly in a blockchain block, each block has:

- **The contents** of the block, for example in bitcoin is it the bitcoin transactions, and the miner incentive reward (currently 25 BTC).

- **A 'header'** which contains the data about the block. In bitcoin, the header includes some technical information about the block, a reference to the previous block, and a fingerprint (hash) of the data contained in this block, among other things. This hash is important for ordering.



Blocks in a chain refer to previous blocks, like page numbers in a book.

Block ordering in a blockchain

Page by page. With books, predictable page numbers make it easy to know the order of the pages. If you ripped out all the pages and shuffled them, it would be easy to put them back into the correct order where the story makes sense.

Block by block. With blockchains, each block references the previous block, not by 'block number', but by the block's fingerprint, which is cleverer than a page number because the fingerprint itself is determined by the contents of the block.

BOOK ORDERING	BLOCK ORDERING
Page 1, 2, 3, 4, 5	Block n58uf0 built on 84n855, Block 90fk5n built on n58uf0, Block 8n6d7j built on 90fk5n.
Implicit that the page builds on the page whose number is one less. eg Page 5 builds on page 4 (5 minus 1).	84n855, n58uf0, 90fk5n, 8n6d7j represent fingerprints or hashes of the blocks.

The reference to previous blocks creates a chain of blocks – a blockchain.

Part 3

DATA STORAGE: *What is a blockchain?*

Internal consistency. By using a fingerprint instead of a timestamp or a numerical sequence, you also get a nice way of validating the data. In any blockchain, you can generate the block fingerprints yourself by using certain algorithms. If the fingerprints are consistent with the data, and the fingerprints join up in a chain, then you can be sure that the blockchain is internally consistent. If anyone wants to meddle with any of the data, they have to regenerate all the fingerprints from that point forwards and the blockchain will look different.

A peek inside a blockchain block: the fingerprints are unique to the block's contents.

This means that if it is difficult or slow to create this fingerprint, then it can also be difficult or slow to re-write a blockchain.

The logic in bitcoin is:

- Make it hard to generate a fingerprint that satisfies the rules of The Bitcoin Blockchain

- Therefore, if someone wants to re-write parts of The Bitcoin Blockchain, it will take them a long time, and they have to catchup with and overtake the rest of the honest network

This is why people say The Bitcoin Blockchain is immutable (can not be changed)*.

**However, blockchains in general are not immutable.*
-> Having said that, the peer-to-peer data sharing mechanism, plus the fingerprinting makes it obvious when a participant tries to alter some data, if you keep track of the fingerprints.

Part 3

DATA DISTRIBUTION: *How is new data communicated?*

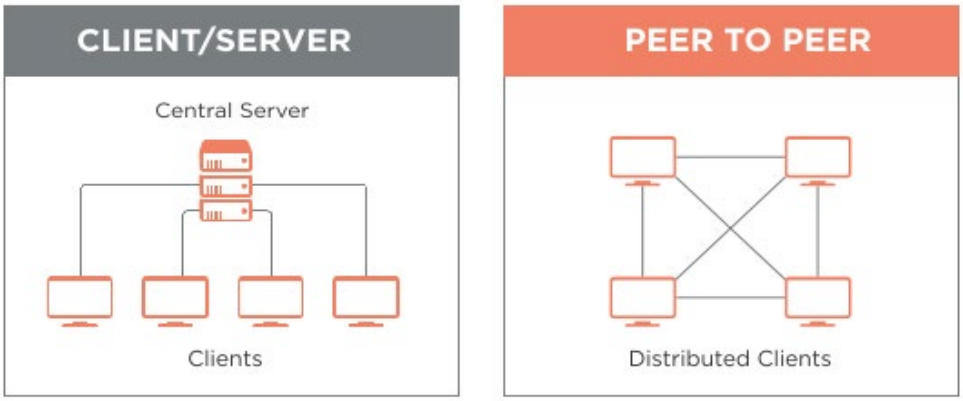
Peer to peer is one way of distributing data in a network. Another way is client-server. You may have heard of peer-to-peer file sharing on the BitTorrent network where files are shared between users, without a central server controlling the data. This is why BitTorrent has remained resilient as a network.

Client-server

In the office environment, often data is held on servers, and wherever you log in, you can access the data. The server holds 100% of the data, and the clients trust that the data is definitive. Most of the internet is client-server where the website is held on the server, and you are the client when you access it. This is very efficient, and a traditional model in computing.

Peer-to-peer

In peer-to-peer models, it's more like a gossip network where each peer has 100% of the data (or as close to it as possible), and updates are shared around. Peer-to-peer is in some ways less efficient than client-server, as data is replicated many times; once per machine, and each change or addition to the data creates a lot of noisy gossip. However each peer is more independent, and can continue operating to some extent if it loses connectivity to the rest of the network. Also peer-to-peer networks are more robust, as there is no central server that can be controlled, so closing down peer-to-peer networks is harder.



The problems with peer-to-peer

With peer-to-peer models, even if all peers are 'trusted', there can be a problem of agreement or consensus – if each peer is updating at different speeds and have slightly different states, how do you determine the "real" or "true" state of the data?

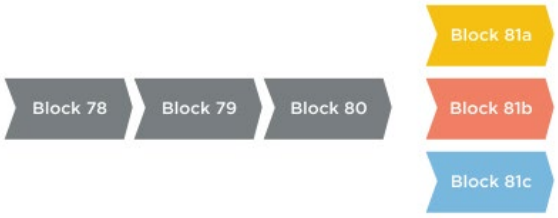
In an 'untrusted' peer-to-peer network where you can't necessarily trust any of the peers, how do you ensure that the system can't easily be corrupted by bad peers?

Part 3

CONSENSUS: *How do you resolve conflicts?*

A common conflict is when multiple miners create blocks at roughly the same time. Because blocks take time to be shared across the network, which one should count as the legit block?

Example. Let's say all the nodes on the network have synchronised their blockchains, and they are all on block number 80. If three miners across the world create 'Block 81' at roughly the same time, which 'Block 81' should be considered valid? Remember that each 'Block 81' will look slightly different: They will certainly contain a different payment address for the 25 BTC block reward; and they may contain a different set transactions. Let's call them 81a, 81b, 81c.



Which block should count as the legit one?

How do you resolve this?

Longest chain rule. In bitcoin, the conflict is resolved by a rule called the "longest chain rule".

In the example above, you would assume that the first 'Block 81' you see is valid. Let's say you see 81a first. You can start building the next block on that, trying to create 82a:

Treat the first block you see as legitimate.

However in a few seconds you may see 81b. If you see this, you keep an eye on it. If later you see 82b, the "longest chain rule" says that you should regard the longer 'b' chain as the valid one (…80, 81b, 82b) and ignore the shorter chain (…80, 81a). So you stop trying to make 82a and instead start trying to make 83b:

Longest chain rule: If you see multiple blocks, treat the longest chain as legitimate.

The "longest chain rule" is the rule that the bitcoin blockchain ecosystem uses to resolve these conflicts which are common in distributed networks.

However, with a more centralised or trusted blockchain network, you can make decisions by using a trusted, or senior validator to arbitrate in these cases.

Part 3

UPGRADES: *How do you change the rules?*

As a network as a whole, you must agree up front what kind of data is valid to be passed around, and what is not. With bitcoin, there are technical rules for transactions (Have you filled in all the required data fields? Is it in the right format? etc), and there are business rules (Are you trying to spend more bitcoins than you have? Are you trying to spend the same bitcoins twice?).

Rules change. As these rules evolve over time, how will the network participants agree on the changes? Will there be a situation where half the network thinks one transaction is valid, and the other half doesn't think so because of differences in logic?

In a private, controlled network where someone has control over upgrades, this is an easy problem to solve: "Everyone must upgrade to the new logic by 31 July".

However in a public, uncontrolled network, it's a more challenging problem.

With bitcoin, there are two parts to upgrades:

1. Suggest the change (BIPs).

First, there is the proposal stage where improvements are proposed, discussed, and written up. A proposal is referred to as a "BIP" – a "Bitcoin Improvement Proposal".

If it gets written into the Bitcoin core software on Github, it can then form part of an upgrade – the next version of "Bitcoin core" which is the most common "reference implementation" of the protocol.

2. Adopt the change (miners).

The upgrade can be downloaded by nodes and block makers (miners) and run, but only if they want to (you could imagine a change which reduces the mining reward from 25 BTC per block to 0 BTC).

If the majority of the network (in bitcoin, the majority is determined by computational power) choose to run a new version of the software, then new-style blocks will be created faster than the minority, and the minority will be forced to switch or become irrelevant in a "blockchain fork". So miners with lots of computational power have a good deal of "say" as to what gets implemented.

Part 3

WRITE ACCESS: *How do you control who can write data?*

In the bitcoin network, theoretically anyone can download or write the software code and start validating transactions and creating blocks. Simply go to *https://bitcoin.org/en/download* and run the "Bitcoin core" software.

Your computer will act as a full node which means:

- Connecting to the bitcoin network
- Downloading the blockchain
- Storing the blockchain
- Listening for transactions
- Validating transactions
- Passing on valid transactions
- Listening for blocks
- Validating blocks
- Passing on valid blocks
- Creating blocks
- Mining' the blocks

The source code to this "Bitcoin core" software is published on Github: *https://github.com/bitcoin/bitcoin*. If you are so inclined, you can check the code and compile and run it yourself instead of downloading the pre-packaged software on *bitcoin.org*.

Permissionless

Note that you don't need to sign up, log in, or apply to join the network. You can just go ahead and join in. Compare this with the SWIFT network, where you can't just download the software and start listening to SWIFT messages. In this way, some call bitcoin 'permissionless' vs SWIFT which would be 'permissioned'.

Permissionless is not the only way

You may want to use blockchain technology in a trusted, private network. You may not want to publish all the rules of what a valid transaction or block looks like. You may want to control how the network rules are changed. It is easier to control a trusted private network than an untrusted, public free-for-all like bitcoin.

For a more in depth look into bitcoin, mining or digital tokens, please see our other free reference papers in the gentle introduction series.

Part 3

DEFENCE: *How do you make it hard for hackers?*

A problem with a permissionless, or open networks is that they can be attacked by anyone. So there needs to be a way of making the network-as-a-whole trustworthy, even if specific actors aren't.

What can and can't malicious attackers do?

A dishonest miner can:

1. Refuse to relay valid transactions to other nodes.
2. Attempt to create blocks that include or exclude specific transactions of his choosing.
3. Attempt to create a 'longer chain' of blocks that make previously accepted blocks become 'orphans' and not part of the main chain.

The attacker can't:

1. Create bitcoins out of thin air.*
2. Steal bitcoins from your account.
3. Make payments on your behalf or pretend to be you.

With transactions, the effect a dishonest miner can have is very limited. If the rest of the network is honest, they will reject any invalid transactions coming from them, and they will hear about valid transactions from other honest nodes, even if they are refusing to pass them on.

With blocks, if the malicious attacker has sufficient block creation power (and this is what it all hinges on), he can delay your transaction by refusing to include it in his blocks. However, your transaction will still be known by other honest nodes as an 'unconfirmed transaction', and they will include it in their blocks.

Worse case, is if the malicious attacker can create a longer chain of blocks than the rest of the network, and invoking the "longest chain rule" to kick out the shorter chains. This lets the attacker **unwind a transaction.**

**Only the attackers version of the ledger will have these transactions. Other nodes will reject this, which is why it is important to confirm a transaction across a number of nodes.*

Part 3

DEFENCE: *How do you make it hard for hackers?*

Here's how you can do it:

1. Create two payments with the same bitcoins: one to an online retailer, the other to yourself (another address you control).

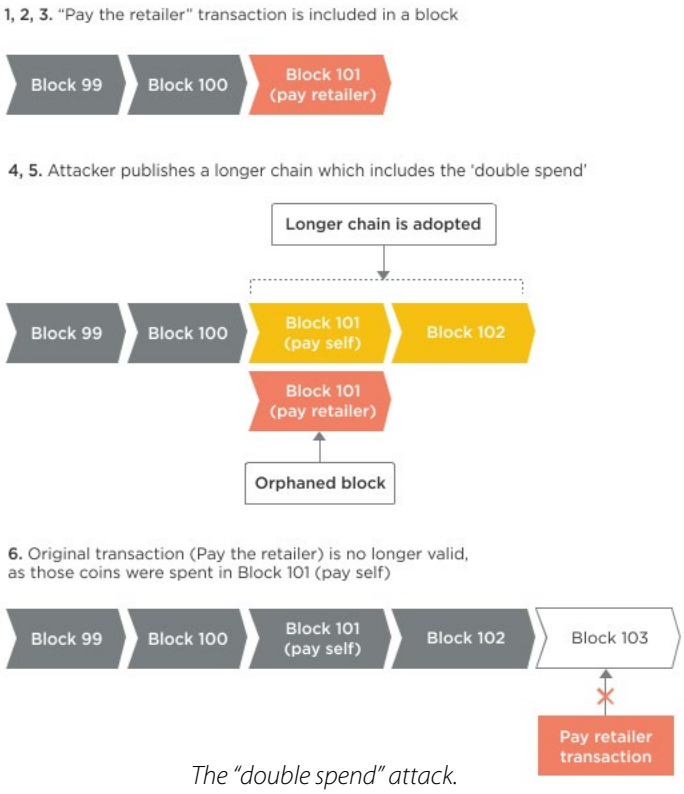
2. Only broadcast the payment that pays the retailer.

3. When the payment gets added in an honest block, the retailer sends you goods.

4. Secretly create a longer chain of blocks which excludes the payment to the retailer, and includes the payment to yourself.

5. Publish the longer chain. If the other nodes are playing by the "longest chain" rule, then they will ignore the honest block with the retailer payment, and continue to build on your longer chain. The honest block is said to be 'orphaned' and does not exist to all intents and purposes.

6. The original payment to the retailer will be deemed invalid by the honest nodes because those bitcoins have already been spent (in your longer chain).



The "double spend" attack.

Part 3

DEFENCE: *How do you make it hard for hackers?*

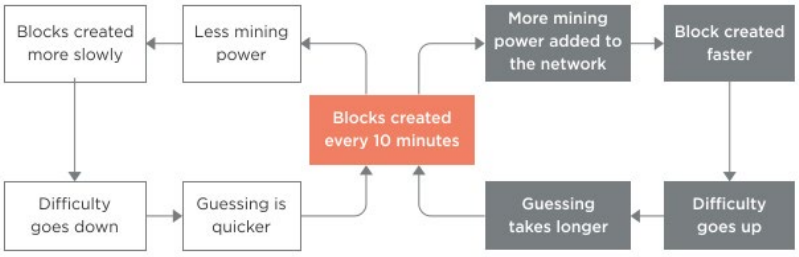
This is called a "double spend" because the same bitcoins were spent twice, but the second one was the one that became part of the eventual blockchain, and the first one eventually gets rejected.

How do you make it hard for dishonest miners to create blocks?

Remember, this is only a problem for ledgers where block-makers aren't trusted.

Essentially you want to make it hard, or expensive for malicious attackers to add blocks. In bitcoin, this is done by making it computationally expensive to add blocks. Computationally expensive means "takes a lot of computer processing power" and translates to financially expensive (as computers need to be bought then run and maintained).

The computation itself is a guessing game where block-makers need to guess a number, which when crunched with the rest of the block data contents, results in a hash / fingerprint that is smaller than a certain number. That number is related to the 'difficulty' of mining which is related to the total network processing power. The more computers joining in to process blocks, the harder it gets, in a self-regulating cycle.



Every 2,016 blocks (roughly every 2 weeks), the bitcoin network adjusts the difficulty of the guessing game based on the speed that the blocks have been created.

This guessing game is called **"Proof of work".** By publishing the block with the fingerprint that is smaller than the target number, you are proving that you did enough guess work to satisfy the network at that point in time.

Part 3

INCENTIVES: *How do you pay validators?*

Transaction and block validation is cheap and fast, unless you choose to make it slow and expensive.

If you control the validators in your own network, or they are trusted, then

- You don't need to make it expensive to add blocks.

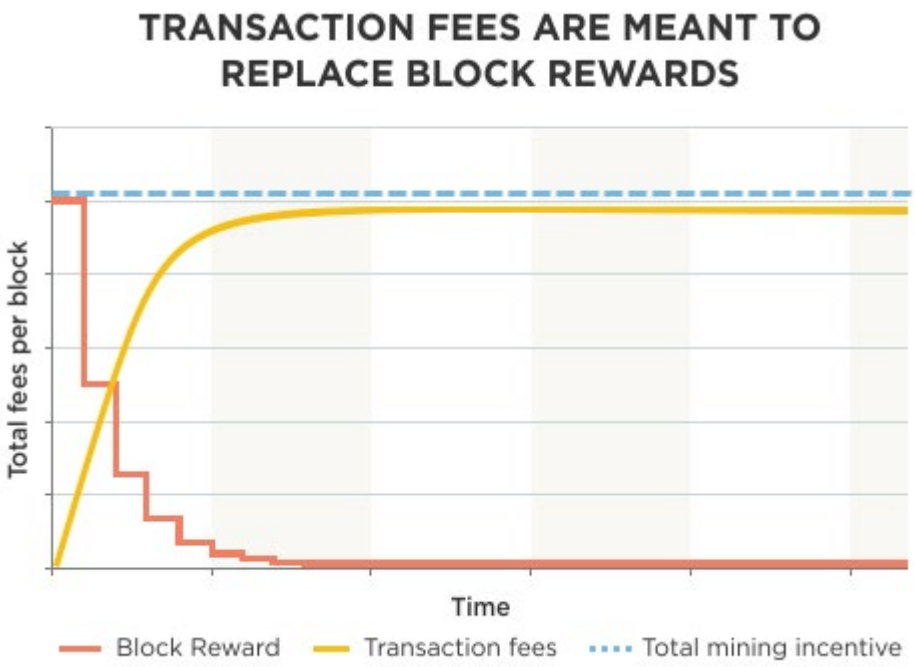
- Therefore you can reduce the need to incentivise them

You can use other methods such as "We'll pay people to run validators" or "People sign a contract to run validators and behave".

Because of bitcoin's 'public' structure, it needs a defence against malicious attackers and so uses "proof of work" to make it computationally difficult to add a block. This has **created a cost** (equipment and running costs) of mining and **therefore a need for incentivisation.**

Just as the price of gold determines how much equipment you can spend on a gold mine, bitcoin's price determines how much mining power is used to secure the network. The higher the price, the more mining there is, and the more a malicious attacker has to spend to bully the network.

The miners do lots of mining, increasing the difficulty and raising the walls against network attacks. They are rewarded in bitcoin according to a schedule, and in time, as the block rewards reduce, transaction fees become the incentive that miners collect.



The idealised situation in Bitcoin where block rewards are replaced by transaction fees.

CONCLUSION

It is useful to understand blockchains in the context of bitcoin, but you should not assume that all blockchain ecosystems need bitcoin mechanisms such as proof of work, longest chain rule, etc. Bitcoin is the first attempt at maintaining a decentralised, public ledger with no formal control or governance. There are significant challenges involved.

On the other hand, private distributed ledgers and blockchains can be deployed to solve other sets of problems. As ever, there are tradeoffs and pros and cons to each solution, and you need to consider these individually for each individual use case.

Puzzled by some of the terms used in these gentle introduction series? Please visit our glossary for a complete terminology breakdown.

www.bravenewcoin.com/bitcoin-basics/glossary/

Acknowledgments

With thanks to David Moskowitz, Tim Swanson and Roberto Capodieci.

About

BNC.
Digital Currency Insights

Brave New Coin is a Data & Research company focused on the exponential Blockchain & Digital Equities industry. We collect, index and report on countless digital assets and their market & industry activities.

Subscribe to our weekly newsletters to keep in the loop with industry news.

Subscribe

www.bravenewcoin.com

contact@bravenewcoin.com

Explore more resources

- Research & insights
- Market-Data
- Developer tools (API's)

Bits on Blocks is a Singapore - based blog, run by Antony Lewis, who focuses on Blockchain Technology. Mr Lewis believes that Blockchain Technology can make the world a better place.

antony@bitsonblocks.net

www.bitsonblocks.net