Blockchains

How crypto-currencies work.

Luke Anderson

luke@lukeanderson.com.au

19th May 2017

University Of Sydney

Overview

Introduction

Suppose we are designing a currency. What properties does it need to have?

Divisibility: A unit of currency should be able to be subdivided into units with equal buying power. Gold is divisible, an iPod is not.

Fungibility: Individual units of currency can be exchanged for each other. Every $10 note is the same, while diamonds are different.

Scarcity: There should be a reasonable restriction on the projected availability of the currency.

Recognisability: It should not be difficult to verify that a piece of currency is genuine.

Limitations of Centralised Digital Currency

Example: Alice has $100 in her PayPal account, and wants to buy some item off Bob for $25.

1. Bob asks for payment from Alice.
2. Alice speaks to the PayPal server and asks to transfer $25 to Bob.
3. Alice tells Bob that the transaction has been processed.
4. Bob checks with the PayPal server and confirms he is now $25 richer.

Limitations of Centralised Digital Currency

Continuing example: Alice has $100 in her PayPal account, and wants to buy some item off Bob for $25.

Advantages:

- Transactions require minimal work from clients.
- Transactions can be reversed, in the case of fraudulent transactions.
- Transactions are secure, and double spending or cheating cannot occur.

Disadvantages:

- The PayPal servers are a single point of failure.
- PayPal can shift or move money at their discretion.
- Alice and Bob can't perform a transaction without the PayPal server.

Bitcoin
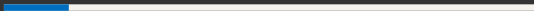
First, there was Bitcoin (BTC)

The first successful "Crypto Currency" payment system, and the first widespread implementation of a Blockchain technology.

- Published 2009 by Satoshi Nakamoto[1]
- Open-source software
- Peer-to-peer
- Worth real money!
  - Now:
    $2,585 AUD / $1,875 USD / $29.8b USD

  See https://blockchain.info/charts
- Trade via Exchanges
  - BTC-e
  - CoinJar
  - Mt. Gox (dead)

---

[1] A pseudonym for an anonymous person or group.

Bitcoin uses public key encryption to secure transactions.

- ○ The *public key* is like a bank account number.
- ○ The *private key* is like a PIN / password.

A *blockchain* takes the place of a central server.

- ○ Transactions are announced to the **peer-to-peer network**.
- ○ **All transactions are visible** to nodes in the network.

Bitcoin *miners* are rewarded to operating the blockchain.

- ○ The creator of the next block is awarded some newly minted Bitcoins.
- ○ The creator of the next block is awarded the transaction fees for transactions processed during the last time period.

A paper published in November 2008 titled *"Bitcoin: A Peer-to-Peer Electronic Cash System"* by the pseudonymous author Satoshi Nakamoto outlined the design of Bitcoin, and the first (open-source) Bitcoin client was released in January 2009.

Nobody knows who Satoshi is, all we know is that they

1. Created the Bitcoin client.
2. Mined many of the first Bitcoins (now worth millions of dollars).
3. Disappeared.

Bitcoin has proven surprisingly strong, both the client program and the protocol itself.

Recall: Currency needs to be *scarce* and *divisible*.

Scarcity:

- ○ There will be at most 21 million Bitcoins (ever).
- ○ They will be released in a predictable fashion.

Divisibility

- ○ Each Bitcoin is divisible 100 million smaller parts, called "Satoshis".
- ○ 1 BTC $= 10^8$ Satoshis.

What is in a Bitcoin wallet?

Create a public/private keypair.

Public: The public key is similar to a bank account number: giving someone this number allows them to transfer Bitcoin to you. This is also called an *address*.

Private: The private key allows you to transfer money *away* from the corresponding address.

Whilst it is possible to use a single address forever, it's very common to use many addresses, even a unique address per transaction.

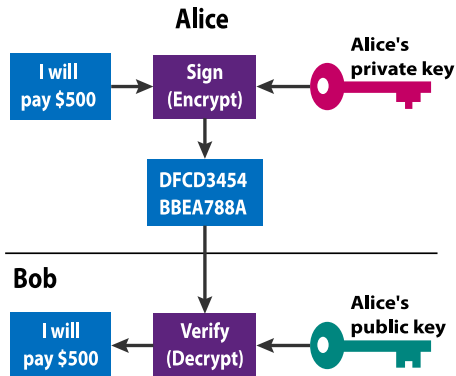| Address (Public key)	Bitcoin Balance
1CkH8epnCee2jSnoYKVf2no8564LygpZcr	1.5027
1Lm9AuUUcazH54qFFW1Rt3V35mNvUCVFb1	17.3723
1D8L2KPG2U8mUqu6seE1GrYCRw2tkCxBHR	0.2
1NbLhL5xGS1YF8LEcXo588EPWSswgsvizb	23.5643

An encryption & decryption key pair are generated.

- ○ **Blockchain address is public key**.
  Anyone can verify your signature with the public key
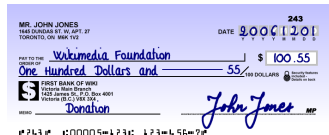
- ○ **Blockchain wallet is private key**
  Only you can sign a transaction with your private key

Alice

| I will pay $500 | → | Sign (Encrypt) | ← | Alice's private key |

↓

DFCD3454 BBEA788A

Bob

| I will pay $500 | ← | Verify (Decrypt) | ← | Alice's public key |

Bitcoin Transactions

Bitcoin
transactions are like cheques.

○ Sender
○ Recipient
○ Amount
○ Signature

We don't accept cheques because
they can't be trusted.

One must *submit it to the bank* (a *central* authority) and wait for it to clear before
accepting you have the money.

In Bitcoin, the sender submits a *transaction to the network* (a *distributed*
authority).

If valid, the network will agree that the receiver is the new owner.

The payee of a transaction must be able to "prove" that the previous owners of the bitcoin did not **double spend** it at any point.

In bitcoin, this is achieved by making everyone in the network aware *of all previous transactions*.

In order to accomplish this without a trusted third party,

1. Transactions must be publicly announced.
2. All participants must agree on a single history for the order of transactions.

The concept of a *blockchain* incorporates a series of interdependant *blocks*, which store a consistent history of information.
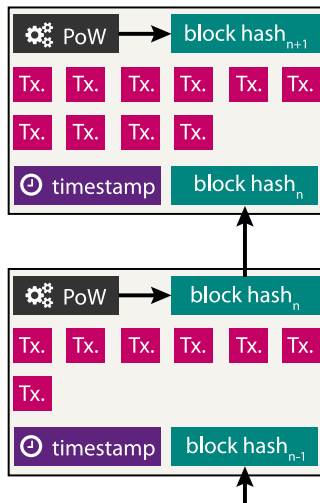
Each block contains:

Tx.: A set of transactions.

block hash: block identifiers, linking them together.

🕐: time the block was "completed".

⚙: Proof of Work – what **miners** do.

Data is **appended only** – with a new block.

○ **The blockchain must definitively record all transactions**
It must show the order in which these transactions occurred, hence it must implement some form of timestamping that is agreed upon by all observers.

○ **The blockchain must be difficult to modify**
Specifically, it should be difficult to modify an existing blockchain such that past transactions may be modified, added, or removed.

The blockchain achieves both these goals by segregating work into chunks, called *blocks*, which are processed on average once every 10 minutes. This time limit, and the difficulty of modifying previous blocks, is accomplished via a *proof-of-work function*.

The blockchain can be thought of as implementing a *timestamp server*, which takes a group of items from timestep n, and combines them into a block B_n.

The hash of current block, H_n, is a function of both the block's contents and the previous hash H_{n-1}.

$$H_n = \mathsf{Hash}(H_{n-1} \parallel B_n)$$

Thus, to modify a previous item, all hashes following the modification must be recalculated.
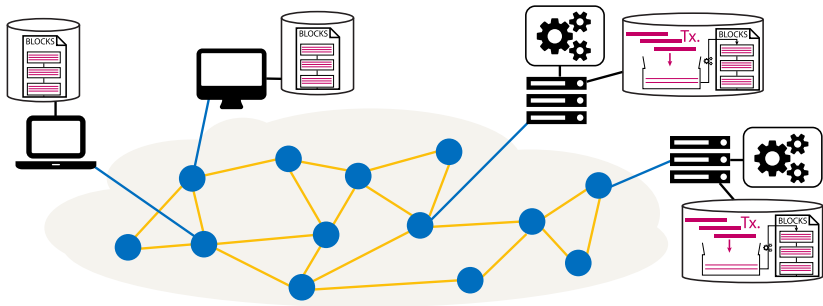
Pure hashing is fast (in fact, most hash algorithms are *designed* to be fast).

The proof-of-work involves making the hashing more time consuming, by including a nonce in the block B_n.

The nonce has to be changed until the hash $\text{Hash}(H_{n-1} \parallel B_n)$ starts with a set number of zero bits.

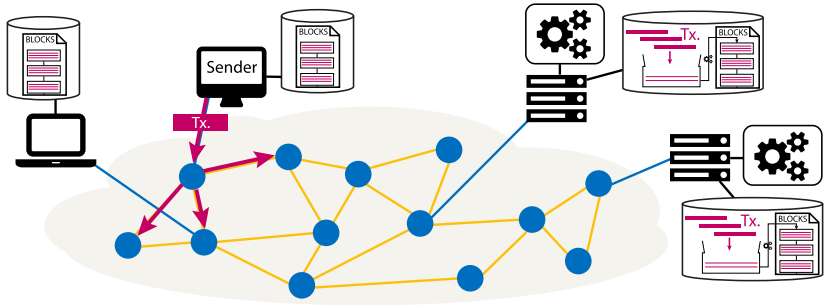| Hash Format	XXXXXXXXXXXXX
Trivial	00XXXXXXXXXXX
Easy	0000XXXXXXXXX
Harder	000000XXXXXXX

All Bitcoin clients are sharing the latest block information via a P2P network.

Some of those clients also *mine* blocks (⚙).
These *miners* are resposible for including transactions (Tx.)
New Tx. are announced on the network by clients and included in new blocks.

A sender authors a transaction Tx. and broadcasts it to the network.
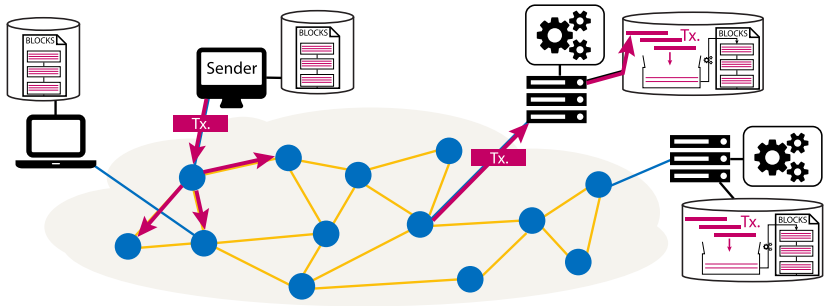
The miners include Tx. in the next block they work on.

One miner **solves** 🏆 a block that includes Tx. .
It announces the new block to the network.
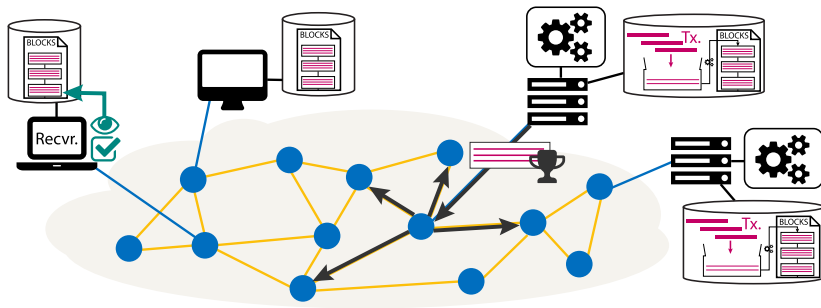
The receiver watches 👁 the blocks to see that Tx. is confirmed.
After a few blocks go by, the receiver accepts ☑ that the funds have cleared.

Ties are resolved when one of the competing heads solves the next block.

The longest chain at any time is the most authoritative.
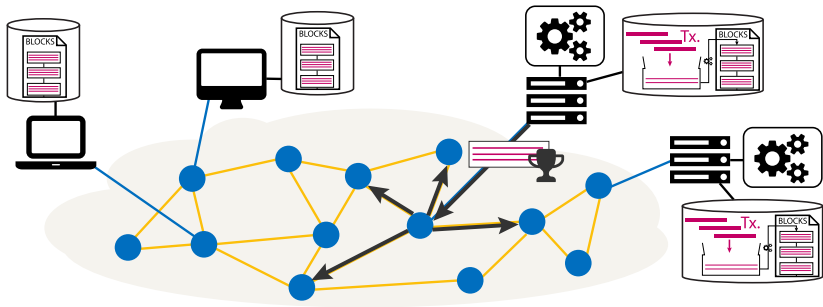
To fork the blockchain, you need to command more than 50% of the computing power in the network.

If you're really unlucky, your transaction will end up in a dead branch, which opens you up to double spending attacks.

Transaction Confirmation
Wait a certain number of blocks before being 'sure' that your transaction will be permanent and irrefutable.
Six confirmed blocks ($\approx$ one hour) is generally considered very safe for Bitcoin.

Incentives for Mining

If you solve the proof-of-work and mine the next block, you receive two significant rewards:

The **mining reward** is essentially "free money in a block" rewarded to the successful miner.

- ○ The only way to create new bitcoin.
- ○ Started at 50 BTC per block, halves approximately every two years.
- ○ Eventually will hit zero.

Transaction fees incentivise miners to include your transaction.

- ○ A successful miner takes the transaction fees associated to each transaction in the block.
- ○ These fees will drive mining in the long run, once the "mining reward" runs dry.

Proof-of-work Difficulty

The SHA256 algorithm outputs 256 bits which should be "approximately a random string of bits", meaning each bit has an equal probability of being a 0 or 1.

- ○ The probability of any hash beginning with 10 zero bits is $2^{-10} \approx 0.1\,\%$.
- ○ The probability of any hash beginning with 50 zero bits is $2^{-50} \approx 10^{-13}\,\%$.

Hash Format	XXXXXXXXXXXXX
Trivial	00XXXXXXXXXXX
Easy	0000XXXXXXXXX
Harder	000000XXXXXXX

Given the speed of hashing across the network (calculated by how fast blocks are mined), bitcoin will adjust how many zeros are required each 2016 blocks ($\approx$ two weeks), to try to make the next 2016 blocks take two weeks.

Bitcoin mining (similar to password cracking) can be done *much* faster on dedicated hardware, rather than general-purpose CPUs.

Ordered by increasing specialisation, mining hardware is categorised as CPU, GPU, FPGA[2], or ASIC[3].

Type of hardware	Millions of Hashes / sec
Multi-Core CPU	35
Average GPU	200
Multi-Card GPU	2000
FPGA	400
$1,000 ASIC	10 000
$30,000 ASIC	14 000 000

An ASIC is entirely custom hardware, and is very expensive to start producing. Both ASICs and FPGAs consume far less power than CPUs or GPUs.

[2]Field Programmable Gate Array
[3]Application Specific Integrated Circuit

"Entire classes of bugs are just missing. Bitcoin has fixed almost all flaws that aren't forced by design."

— Dan Kaminksy

The bitcoin protocol is **open source**, and has been since the beginning.

Bitcoin's security resides in three places:

○ The strength of public-key crypto (ECDSA) to protect accounts.
○ The strength of SHA256 in the proof-of-work function.
○ General secure programming.

Even if one of these were broken in the future, Bitcoin is made to be upgradeable.

Improper verification. Verification wasn't properly done on transactions before they entered the block chain and less than a week after discovery a fraudulent transaction resulted in 184 billion fake Bitcoins being created (reverted by the community).

Blockchain Forks. The blockchain temporarily forked into two independent chains due to a major software bug. The new Bitcoin client produced a transaction that wasn't accepted by the older client, splitting the blockchain and producing the first real world examples of "double spending".

Control of the Network. If an attacker had more than 50% of the computing power of the network, the attacker could perform double spending and also reject other people's transactions from receiving confirmations.

Blockchain Security

Hardware is hard.

This makes it useful for key storage because breaking into hardware is difficult.

Unfortunately, developing and deploying hardware is also very difficult.

Some challenges:

- Hardware still depends on software.
  - Need more sophisticated hardware to counteract
    - case study: Trezor

- Can give users a false sense of security
  - e.g. falling victim to card skimming

- Carrying hardware sucks
  - How many cards are in your wallet?
  - How many RSA tokens on your keychain?
  - **Solution:** standardisation & cooperation
    - case study: Yubikey

- Developing hardware is expensive

Hardware solutions have a ways to go.

Need to consolidate solutions while allowing flexibility and upgradability.

Need to:

- ○ Establish public and academic trust
  - ○ Open-source is essential
  - ○ Security through obscurity $\neq$ security
- ○ Prevent firmware updates to prevent key theft
- ○ Allow updates to enable future-proofing

Distributing Responsibility

Since we can't trust people to look after their own keys, we need ways to spread the risk to *multiple* people.

It's harder to hack all 5 individuals, or have 5 individuals all loose their keys, than one.

Copying the key to 5 people reduces likelihood of losing key, but increases likelihood of it being stolen or misued.

How do we distribute the responsibility of a key?

○ Need to hack 2 or 3 people.
○ If one person loses their key, the others can "reset" or "reissue" their key.

Distributing Responsibility: Multisigs

Multisignature is a technique to enforce complex key requirements.

In a multisig scenario, a group of n people have their own independent keys which they self-manage.

A piece of software places business-logic requirements on which keys must be present in order to perform an action.

Multisig in Bitcoin

Bitcoin has a built-in scripting language that allows software authors to determine rules.

For example:
"At least 4 of the 6 authorised keys must sign a transaction in order for the funds to be released."

Distributing Responsibility: Key Splitting

Multisigs are good when our protocol supports it, however there are many applications where multisig doesn't make sense.

PERSONAL CRYPTO FORTUNE

"How do I give my family access to my diverse cryptocurrency fortune if I die?"

Normal circumstances:

○ Only I have access.
○ I have full control myself.

Extenuating circumstances:

○ Many trusted 3rd parties must come together
○ A single rogue 3rd party cannot take control

Could be done with multisigs, but need a different implementation for each cryptocurrency.

Shamir's Secret Sharing is an algorithm for dividing a secret into m pieces, where only n of them are required to reconstruct the original secret.
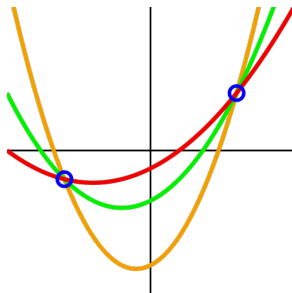
A polonomial of degree n can be uniquely defined by plotting $n + 1$ points on that polynomial.

Example

$$y = ax^2 + bx + c$$

Infinite possibilities for the coefficients with only 2 points.

3 points, and a, b, and c are uniquely determined.

So keeping crypto is hard, let someone else do it for you!

A lot of people store their cryptocurrency on exchanges.

Not actually *that* bad when you consider security vs. usability.

- ○ Authenticated with:
    - ○ Password
    - ○ 2FA
    - ○ E-mail address
- ○ *Very* famililar
- ○ *Very* easy to use

But people don't understand the risks.
Many of these companies have incompetent tech people, and security is *hard*.

Getting hacked is inevitable.

How you respond to a hack affects how it impacts you.

Blockchains are thought to be anonymous.
In general, they're not. They're **pseudonymous**.

Like all security, if your attacker is large enough, you cannot hide.

If someone can monitor the vast majority of the internet, bitcoin's
anonymity is out the window.

Ethereum, however, is not.
It's P2P network protocol has built-in encryption with forward-secrecy.

GAME THEORY

Game theory is:

"the study of mathematical models of conflict and cooperation between intelligent rational decision-makers."

Many attacks and security considerations in decentralised systems are solved using game theoretic solutions.

Blockchain systems are often said to be "trustless" or that they "remove the need to trust a 3rd party".

This is straight up **false**.

Blockchains do not remove trust, they *move* trust

Instead of trusting a central entity (e.g. a bank) to perform some action with integrity, the users are trusting a whole network of computers to perform some action with integrity.

A network is a lot more complicated than a central authority, and has many points of failure.

There are many other kinds of attacks on the network, such as the Denial-of-Service attacks seen on the Ethereum network.

By choosing incentive structures carefully (again, game theoretic), these can be avoided.

Ethereum EVM

The Ethereum EVM opcode: EXTCODESIZE costs a normal amount of 'gas', but requires significant resources to read from disks.

By spamming the network with transactions that called this 50,000 times per block, the network was brought to a halt in September 2016.

Since humans rely on other humans to make decisions, key personalities in the community can have a profound effect on decentralised systems.

For example, Vitalik Buterin has a *massive* influence on the price of ETH.

DAO Hack

When the DAO hack was ongoing, the price of DAO tokens crashed.

The moment Vitalik posted publically that the network would hard-fork and refund people's money, the price almost returned to normal.

Blockchain networks are about *human consensus*, not about *immutable code*.

Beyond Bitcoin

Namecoin was the first "altcoin" and extends Bitcoin functionality by acting as a distributed DNS system.

A transaction in Namecoin extends Bitcoin's transactions:

○ Name Update

Users can register ".bit" domains (e.g. usyd.bit) by submitting a transaction with the "Name Update" command.

Ethereum fully extends the concept of transactions, by creating "smart contracts".

A contract is like an independent entity that is governed by it's (full-turing) code. It has data storage, it's own currency balance, and can be interacted with via transactions.

Now we can easily run arbitrary software on a blockchain without creating a whole new chain.

Ethereum has an internal currency, "ether", which is burned as "gas" when executing contracts.
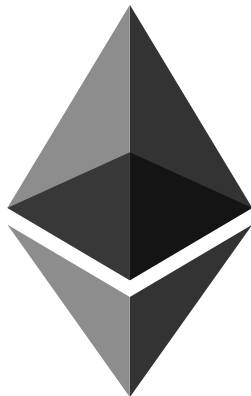
Smart Contracts

A **smart contract** is a program on the blockchain.

Like a regular program, a smart contract has:

○ Code
○ Inputs
○ Outputs
○ Data storage

Unlike a regular program, a smart contract is executed by *all* nodes in the network.

Anyone can send a message (input) to a smart contract, causing all nodes in the network to agree on the state change caused as a result of the message and code.

A smart contract:

○ is fully defined by it's initial code.
  (code is public)
○ cannot be modified.
  (unless it's code explicitly permits)
○ can receive input from anyone on the network.

Since a smart contract could fully encapsulate business logic, they can be used to create:

decentralized autonomous organisations (DAOs).

The premise[4] is simple:

○ Initial balance: 100 ether.
○ Send at least 1 ether to participate
○ If 24 hours elapse with no deposit:
   The last participant receives all Ether sent by all participants.

Before Ethereum, this would be a scam.
The entity running this could easily run away with all the money.

But now, our entity is a smart contract. It is governed by it's code, and controlled by nobody. It's completely fair and trustworthy.

http://lottopollo.com/

---

[4] https://www.reddit.com/r/ethereum/comments/3ivfyn/lotto_pollo_lottery_chicken/