Department of Computer Science and Engineering

CSC405A: Seminar

Architecture & Application of Blockchain

Name:	Subhendu Maji	Name:	Tanishq R Porwar
Reg No:	18ETCS002121	Reg No:	18ETCS002131
Dept:	CSE	Dept:	CSE
Batch:	2018	Batch:	2018

Outline

- Introduction
- Motivation for choosing the Topic
- Presentation of Topic
- Student’s opinion/perspective
- Relevance to society
- Conclusions
- References

History

- During late 2008, and the global financial crisis was causing shock waves around the world.
- Anger at the worldwide banking industry, governments and other centralized authorities has reached fever pitch.
    - Enter a mysterious figure named **“Satoshi Nakamoto”**, whose real identity continues to remain shrouded in mystery to this day.
- Satoshi authors and releases a white paper titled **Bitcoin: A Peer-to-Peer Electronic Cash System**.
- The paper shared the workings for a new digital currency system that didn’t rely on banks to facilitate transactions or governments to create and disseminate the currency.

Motivation

- When it comes to transacting money or anything of value, people and businesses have relied heavily on intermediaries
  - such as banks and governments to ensure trust and certainty.
- Middlemen perform a range of critical tasks that help build trust into the transactional process.
- Things like payment authentication & record keeping. The need for intermediaries is especially acute when making a digital transaction.

Double spending Problem

- If you can send someone $100 online, yet still, have that original $100 under your name.
  - That would mean you could just keep spending that $100 as many times as you wanted.
  - The money would become meaningless.
  - This problem doesn't exist in the physical world. After a person spends physical currency like US dollars, they no longer have that cash (the actual notes) in their possession. They can't, therefore, spend the same money over and over.
- For example, if you spend $100, banks ensure that your account balance decreases by $100 and the account of the person or organization you transacted with increases by $100. No double spending can occur

Double spending Problem

Problems:

- Merchants need information about customers to build trust as transactions are reversible
- Minimum transactions cost rise as banks can’t avoid mediating disputes

What is a blockchain?

- A blockchain is a type of distributed ledger or decentralized database that keeps continuously updated records of digital transactions (who owns what).
- The original Bitcoin blockchain is designed as a write once read only database where records can only ever be added, not edited or deleted.
- Rather than having a central administrator like a traditional database, (banks, governments), a blockchain has a network of replicated databases, synchronized via the internet and visible to anyone within the network.

How does this decentralized network (the Bitcoin blockchain) overcome the double spending problem?

- It does this by publicly announcing all transactions to the network.

- As Satoshi states:
  - *“The only way to confirm the absence of a transaction is to be aware of all transactions. In the mint based model, the mint was aware of all transactions and decided which arrived first. To accomplish this without a trusted party, transactions must be publicly announced.”*

Cryptographic hash function

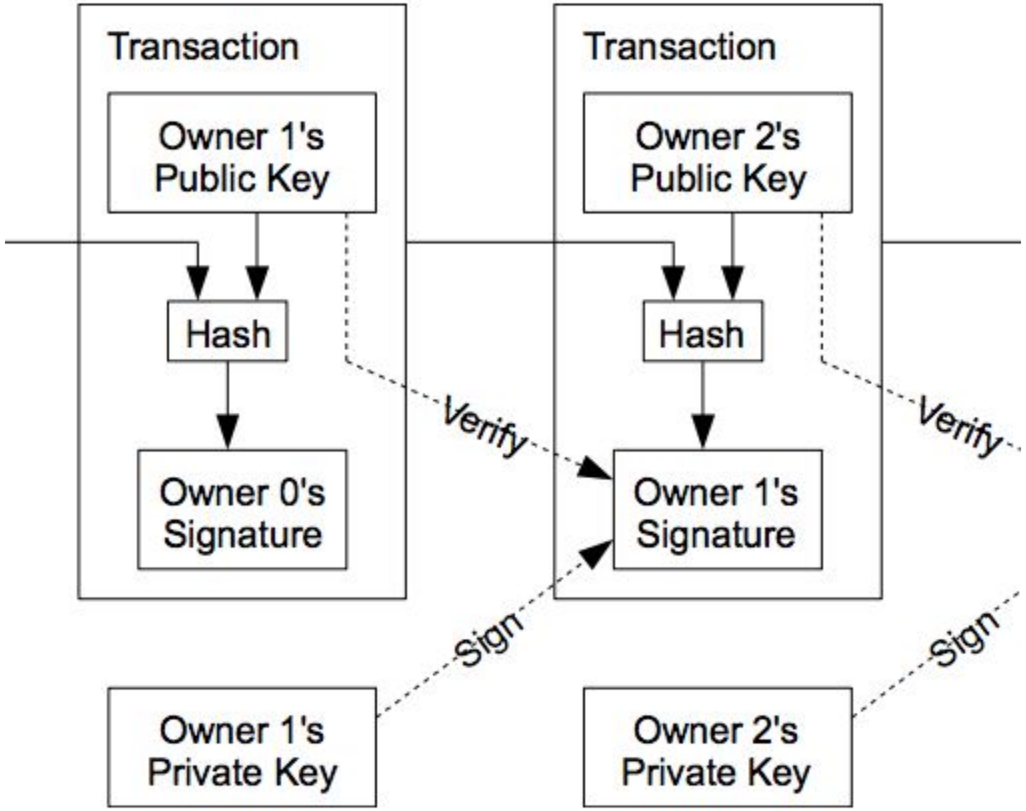
- **Cryptographic hash function**:
  - we can picture a *hash function* as a blackbox that takes a string as input (such as “Hello Bob”) and returns a fixed size arbitrary string (such as “98b0f4b363af4aceb81bc42fd81117e1”).

For a hash function to be usable cryptographically it must have certain properties:

- Same input always returns same output
- It’s quick to compute
- we can’t reverse engineer the “98b0f4b363af4aceb81bc42fd81117e1” that comes from “Hello Bob” without brute force (trial and error)
- A small change in the input will change the output a lot
- It’s unfeasible that two inputs generate the same output

Asymmetric cryptography: it will allow us to communicate through an insecure channel. A typical use case is when Bob wants to send a message to Alice which only she can read. Another use case is to be able to verify who the sender of a certain message is. This is the case that we are interested in

The Chain


Transaction
Owner 1's Public Key
Hash
Owner 0's Signature
Transaction
Owner 2's Public Key
Hash
Owner 1's Signature
Verify
Verify
Sign
Sign
Owner 1's Private Key
Owner 2's Private Key

Now question we need to ask

- what happens if members of the network use different transaction timelines?
- Members are spread around the world so won’t people be able to double spend their bitcoins?
- How do participants in the Bitcoin network agree on a single history of the order in which transactions were received?

Timestamp server

- To avoid these issues, members of the network agree to a single transaction timeline and process transactions according to their timestamp
- Even though the majority of the network agree to run on a single timeline, for a decentralized system like Bitcoin to operate without any central intermediary, there needs to be a way for the network to agree about which order transactions are generated in. That means each transaction needs to get stamped with a precise time on it
- The timestamp server is a piece of software that timestamps transactions when they occur. It takes a small section of the transaction data and digitally timestamps it to create a hash.

What happens after the hash is created?

- The timestamped hash is made publicly available for everyone in the network to view.
- The network processes each transaction in order of their respective timestamped hash.
- The hash serves as a complex computer problem that needs to be solved by miners before a transaction can be added to the blockchain for eternity.
- Each time stamp includes the previous transaction timestamp thus forming a chain of transactions aka a blockchain.

Mining

- Proof of Work aka mining is performed to facilitate transactions on the blockchain and discourage bad actors from spamming the network by sending out fraudulent or illegitimate transactions.
- e.g.

  An example could be reducing email SPAM: I give you a challenge to be resolved before sending the email which will cost you some computing power (i.e. electricity hence money). It will be small enough so that sending one email is cheap but costly enough to avoid a spammer send many.
- It involves miners (members in the network with high levels of computing power) to prove that a specified amount work has been completed.

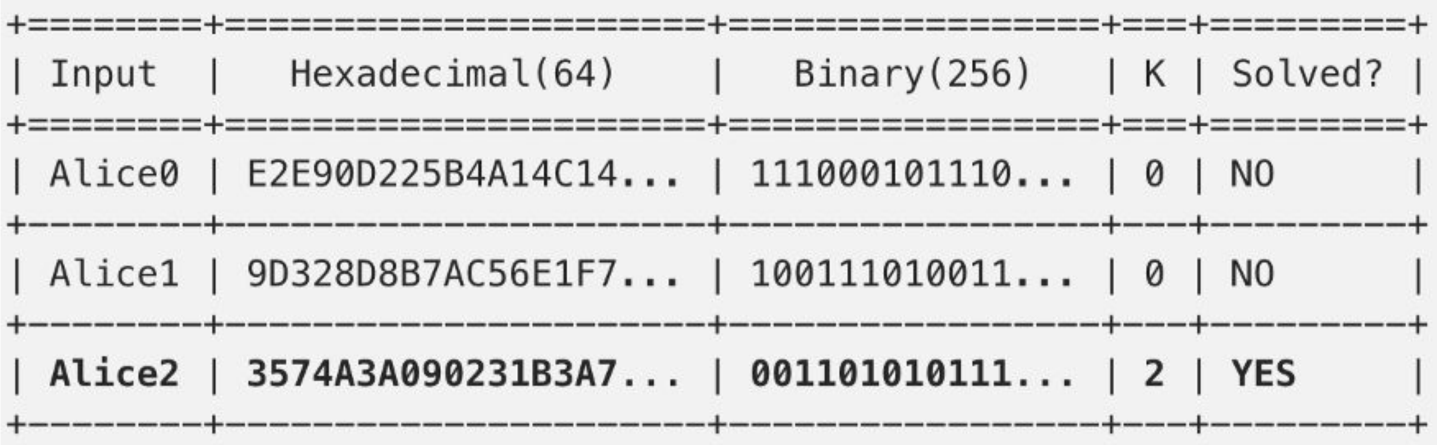
What if someone hacks ? …………well they can try

HASHCASH

Let’s say the content of the e-mail you are sending is “Alice”.

We will hash the message before we send it. A challenge I could give you is to find a number which when appended to the content (i.e. “Alice24”) will produce a binary output with certain amount of 0’s in the beginning.

So let’s say *k* is the number of 0’s we are asking for. Given the input “Alice” and a challenge *k=2*:



```
+========+======================+=================+===+=========+
| Input  |   Hexadecimal(64)    |   Binary(256)   | K | Solved? |
+========+======================+=================+===+=========+
| Alice0 | E2E90D225B4A14C14... | 111000101110... | 0 | NO      |
+--------+----------------------+-----------------+---+---------+
| Alice1 | 9D328D8B7AC56E1F7... | 100111010011... | 0 | NO      |
+--------+----------------------+-----------------+---+---------+
| Alice2 | 3574A3A090231B3A7... | 001101010111... | 2 | YES     |
+--------+----------------------+-----------------+---+---------+
```

We have found a solution by calculating the hash 3 times. Everyone can calculate the hash just once and verify that what we solved the challenge correctly. We could make the challenge much more difficult by requiring more 0’s to appear on the output (k=20 requires on average 1 million tries, k=30, billion tries).

Proof of Work

- We need to find a way to give incentives to distributed machines to find consensus (vote) on what is the current state and order of the transactions within the blocks that form the blockchain. If there was no cost associated in generating each block it would be costless for anyone to manipulate it (such as in the SPAM prevention example) and hence the system would be hackable by definition.
- By applying a variation of Hashcash PoW we can accomplish this:
- Instead of using “Alice” as we did in the last example, what bitcoin does is applying the hash function on the headers of the block (which include the previous block hash, the timestamp and the nonce among others). The nonce is the number we have to increment until we satisfy the 0s challenge we are solving just as in our example.
- An immediate question arises: The nonce is a number defined by 32bits (allows ~4 billion tries)
- But isn’t that a quite easy challenge to solve for a miner with current technology? That’s why there is an extraNonce in every block which has to be modified by the miner whenever the 4 billion hashes have been calculated with no success. A modification on the extraNonce changes the Merkle tree root so we can start calculating hashes from 0 again.

Proof of Work

- This is how a block is mined. As there are multiple nodes within the network they have to do decision making through consensus. What a miner considers truth is validated by the PoW it has performed rather than something like an IP address, which could be tricked by a single entity
- Decision making is achieved by the longest chain:
- If the majority of CPU is held by honest nodes, their chain will outpace attackers: the probability of changing a past block and catching up diminishes exponentially by every block that is added as the attacker would have to redo the work of that block, all blocks after it and surpass the current nodes work.

The network

The steps involved are as follows :

1. New transactions are broadcast to all computers (nodes) in the network.
2. Each node collects new transactions into a block of transactions.
3. Each node works on finding a difficult proof-of-work for its block.
4. When a node solves the mathematical problem (proof-of-work), it broadcasts the block to all nodes.
5. The network nodes only accept the new block if all transactions in it are valid and not already spent.
6. Nodes then move on and start creating the next block in the chain.
7. Repeat above steps.
8. If two nodes broadcast different versions of the next block simultaneously, the network nodes consider the longest chain to be correct and will keep working on extending it.
   a. Any nodes that are switched off and fail to receive a new block will be updated when they connect back to the network.

Incentive

- Bitcoin mining is an expensive and time-consuming task.
- To incentivize members to support the network a reward is given in the form of bitcoins.
- The first transaction in a block creates a new coin which is owned by the person (node/miner) who solved the puzzle and subsequently created that particular block.
- This adds an incentive for nodes to support the network, and provides a way to initially distribute coins into circulation, since there is no central authority to issue them.

Reclaiming disk space

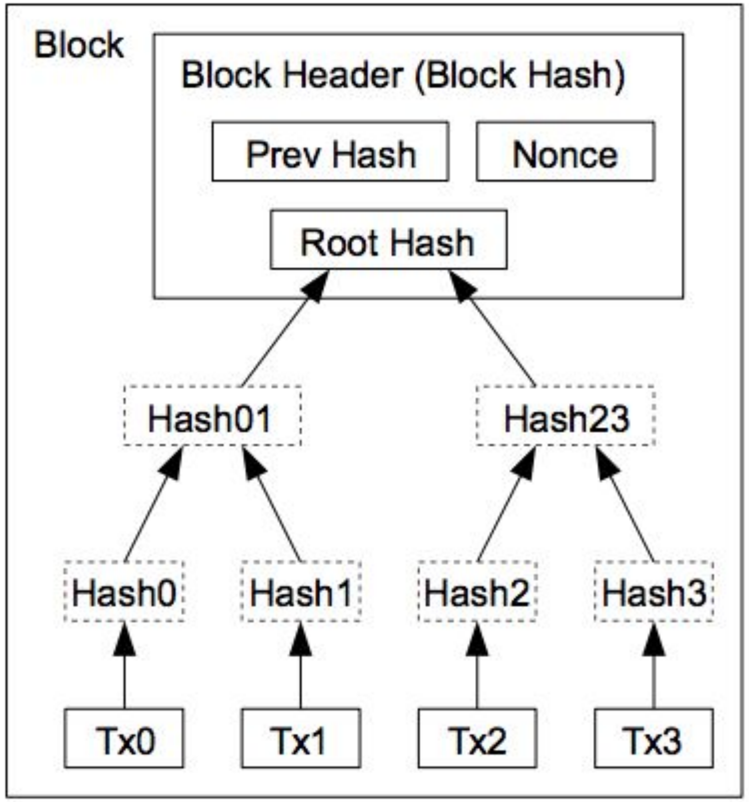
- the average transaction size since 2014 is around 520 Bytes.
- There are around 277 Million transactions for that period.
- Excluding the size of the blocks it would mean that storing all these transactions would occupy ~144 GB. In reality they occupy ~150GB.

The block header with no transactions occupies ~80 Bytes. The number of blocks for the same period is ~240,000 which occupies a total of 19MB. Such size sounds much more competitive.

Merkle tree

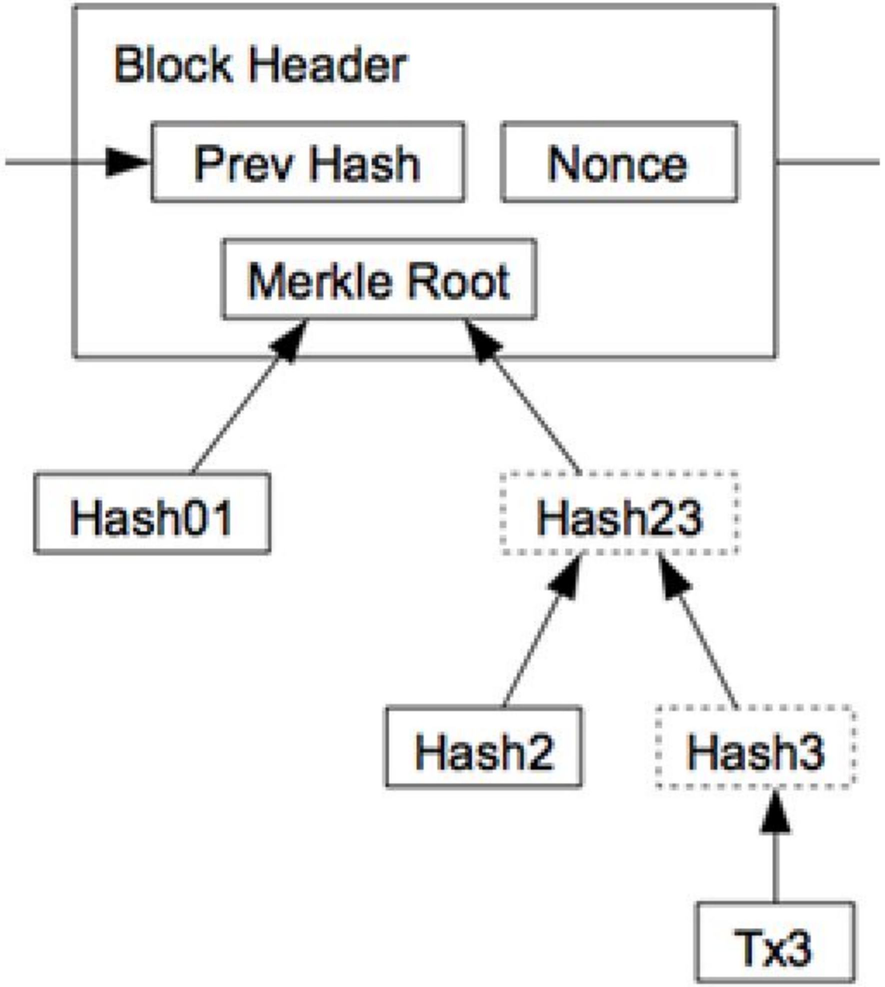
- The leaf nodes (Hash0, Hash1, …) are a hash function applied to some data (in this case Tx0, Tx1, etc…).
- Upper nodes (Hash01, Hash23) are just a hash function applied to the concatenation of their respective children hashes.
- The *Root Hash* is the upper-most hash which is included in the block header and ensures what transactions are present.

- Why don't we just concatenate all hashes one after the other and produce the root hash out of that string?
  - ○ Well, if we wanted to validate if a transaction is part of a block which has many transactions, a Merkle Tree would allow us to do it in logarithmic cost compared to getting all hashes.



Simplified Payment Verification

- Everyone don’t have to be a miner that helps verify transactions to make Bitcoin transactions.
- It’s also possible to just send and receive bitcoins with a simple Bitcoin wallet.
- Most members of the Bitcoin network around the world do not operate full payment verification nodes and don’t have massive supercomputing power at their fingertips.



Student’s opinion/perspective

- With many practical applications for the technology already being implemented and explored, blockchain is finally making a name for itself, in no small part because of bitcoin and cryptocurrency.
- As a buzzword on the tongue of every investor in the nation, blockchain stands to make business and government operations more accurate, efficient, secure, and cheap, with fewer middlemen.
- it’s no longer a question of if legacy companies will catch on to the technology—it’s a question of when.
- Today, we see a proliferation of NFTs and the tokenization of assets. The next decades will prove to be an important period of growth for blockchain.

Relevance to society

Today, there are more than 10,000 other cryptocurrency systems running on blockchain. But it turns out that blockchain is actually a reliable way of storing data about other types of transactions as well.

- Banking and Finance
- Supply Chains
- Smart Contracts
- Healthcare

Conslusions

Here are the key takeaways :

- To overcome the double spending problem which results in reliance on intermediaries and a whole new set of problems (inability to make non-reversible transactions, increased costs, etc.) Satoshi proposes a new electronic payment system that relies on complex computer encryption (cryptography) instead of the trust generated by intermediaries.
- A blockchain is a type of distributed ledger or decentralized database that keeps continuously updated records of digital transactions (who owns what). It is the underlying technology that enables Bitcoin to operate.
- Instead of relying on centralized intermediaries to provide security and privacy, Bitcoin transactions use cryptography. Transaction information can't be linked to any identify because it is encrypted. Members of the network only see a random bunch of letters and numbers.
- For a decentralized system like Bitcoin to operate without any central intermediary, there needs to be a way for the network to agree about which order transactions are generated in (to prevent double spending) and which transaction records are valid (to deter any abuse of service like denial of service attacks and spamming).
- Proof of Work aka mining is performed to facilitate transactions on the blockchain and prevent abuse of service attacks. It involves miners (members in the network with high levels of computing power) to prove that a specified amount work has been completed.
- To incentivize members to support the network and carry out the expensive and time- consuming task aka mining, a reward is given in the form of bitcoins.
- To maximize disk space and keep the entire history of the Bitcoin blockchain intact, the Bitcoin network keeps a trace or root of transaction data.
- You don't have to be a miner that helps verify transactions to be involved in the Bitcoin network. It's also possible to send and receive bitcoins with a simple Bitcoin wallet

References

- https://bitcoin.org/bitcoin.pdf
- https://zerocap.com/the-bitcoin-whitepaper-summary/
- https://static1.squarespace.com/static/567bb4f069a91a95348fa0b2/t/5cd27c8bb208fcb3a45d2196/1557298317565/Intrepid+Ventures+Bitcoin+White+Paper+Made+Simple.pdf
- https://www.investopedia.com/terms/c/consensus-mechanism-cryptocurrency.asp