CSE5095-010: Blockchain Technology

Lecture 6

Ghada Almashaqbeh
UConn - Fall 2020

Outline

- Ethereum.
  - Work model.
  - Account types.
  - Transaction processing.
  - Mining and consensus.
  - State machine (or blockchain).

Motivation

- Bitcoin has several limitations, among them:
  - No user accounts, tracking a currency balance is simply tracking a chain of transactions.
    - I.e., the UTXO model.
  - Limited scripting language, no-Turing complete that supports a small set of instructions.
- This motivated creating more flexible systems that allow users to ask the mining network to implement any program they wish.
  - Such programs are known as DApps or smart contracts.
- This lead to smart contract-enabled blockchain systems.
- We will explore the first and most widely used smart contract-enabled system; ***Ethereum***.

Ethereum’s History

- Proposed by Vitalik Buterin in 2013 and went live in 2015.
- Supports a permissionless, proof-of-work based blockchain.
- Its native token called Ether (ETH).
    - Intended to be a utility token to pay for computation.
- Its ability to execute user programs, aka smart contracts, made it a very attractive option for numerous applications.
    - No need to spin a new system, use Ethereum’s infrastructure for even creating new cryptocurrencies!
- Because of the DAO incident, Ethereum experienced a hard fork in 2016.
    - Led to two separate systems; Ethereum and Ethereum Classic.

The Big Picture

- Build a general-purpose blockchain.
- Allow users to instruct the miners to run user-customized functionalities and computations.
    - Supports a Turing-complete scripting language to write smart contracts.
- View the miners as a global virtual computer to execute smart contracts.
    - Users can deploy smart contracts on the blockchain.
    - Users can invoke functions in a smart contract, and the miners execute such function calls.
    - Function calls are packaged as transactions.
- The global computer is called Ethereum virtual machine (or EVM).

EVM

- Used to implement smart contracts, or distributed applications (DApp), deployed by clients.
- The code of these smart contracts are recorded publicly on the blockchain.
- Each instruction in any contract is executed by every miner in the network.
    - Those who hear the function calls submitted by clients.
- Changes in the smart contract state (i.e., memory/variable values/etc.) after these function calls are also recorded on the blockchain.
    - Can be verified by anyone.

ETH GAS STATION

Computation Costs Money I

- How about DoS attacks?
  - E.g., deploy programs with infinite loops that can stay active in the EVM forever.
- Miners charge a fee for each instruction they execute.
  - These include arithmetic computations, data access, flow control, etc.
- This fee is called gas.
  - Each instruction type has a designated price in gas units.
  - A transaction issuer has to provide the suitable fee in order for the miners to implement the requested operations.

ETH GAS STATION

Computation Costs Money II

- Gas is purchased using Ether.
  - So a transaction issuer sends Ether that is used to buy gas based on the gas price it is willing to pay.
- Gas price is not fixed, it is miner dependent.
  - Miners announce their gas prices, check https://ethgasstation.info/
  - The higher the gas price a tracation issuer is willing to pay, the larger the number of miners willing to process a transaction.
- A miner computes the number of gas units, then charge the issuer in Ether.
- Any extra fees are refunded to the issuer.

Main Components

- Similar to other permissioned, public cryptocurrencies, Ethereum's main components are:
  - P2P network.
  - miners/clients.
  - Transactions.
  - Mining and consensus rules.
  - Blockchain.
  - Economic security.
- Different from UTXO-based cryptocurrencies, Ethereum has:
  - Account-based model.
  - State machine to track system changes.

Account Types

- Externally Owned Accounts (EOAs):
  - Associated with a private/public key pair.
  - The user who owns this key has full control of the currency in this account.
- Contract account:
  - Associated with a smart contract code.
  - Does not have a private key, it is owned and controlled by the logic of the code.
- Both account types have addresses.
  - EOA: derived from its public key (hash then take least 20 bytes).
  - Contract accounts: derived from the creator's address (the user who deployed the smart contract) and his/her account nonce.
- One needs an EOA to create a contract account.

Transactions I

- Transactions can be initiated by EOAs.
- To prevent replay attacks, each EAO has a counter that increments after each issued transaction (usually called a nonce).
- The notion of accounts and contracts make the structure of a transaction much different than in Bitcoin.
  - No need to reference other transactions as input.
  - Just reference the account that the sender owns (this will be used to deduct gas fees and transferred currency).
- Transactions can be:
  - Standard currency transfer.
  - Deploying contracts.
  - Function calls.

Transactions II

- The destination of a transaction can be:
  - EOA:
    - Usually used for currency transfer.
    - Leads to updating the balance of the sender and receiver.
  - Contract:
    - Causes a code in the contract to be executed using the data in the transaction as input.
    - Updates the contract state on the blockchain (and EOAs if any).
  - The address zero.
    - Used when deploying a contract, or known as registering it on the blockchain.
    - The payload is the compiled bytecode of the contract.

Transactions III

- The fields of a transaction are:
  - Nonce (or a sequence number).
  - Gas price (gas unit price the issuer is willing to pay).
  - Gas limit (total amount of gas the issuer is willing to buy for the transaction).
  - Recipient.
  - Value (amount of currency to be transferred including the fees).
  - Data (function inputs, etc.).
  - Signature.
- Processing a transaction means validating its format, fees, updating account status (if any), execute a function call and update a contract state (if any), and refund rest of fees (if any).
- Each transaction is recorded on the blockchain.

Ethereum’s Blockchain Explorer

- Visit https://etherchain.org/
  - How long does it take miners to generate a new block on average?
  - Does the shorter block time increases the transactions per second (check the TPS value)? Why is that?
  - Are 6 blocks enough to confirm a transaction in Ethereum as in Bitcoin? Why?
  - Open the statistics tap. Are there mining pools in Ethereum blockchain?

Mining I

- Currently it is proof-of-work based, with a slightly different version than the one used in Bitcoin.
    - Called Ethash.
- Ethash is a memory-bound algorithm instead of computation-bound.
    - To be an ASIC-resistant algorithm that is controlled by memory access cost instead of computation cost.
- At a basic level, each miner generates a pseudorandom dataset, called a DAG (Direct Acyclic Graph), that is expanded every 30K blocks.
    - The seed is derived from the current length of the blockchain.
    - All miners, who have the same blockchain view, will generate the same DAG.
    - Initial size was 1 GB, now it is around 3.5GB.

Mining II

- The candidate block header and the nonce (a guess for the hash puzzle solution) are used to select a random subset of the DAG.
- The DAG subset, the header, and the nonce are all hashed together.
- If the output meets the network difficulty, a valid solution has been found.
- Other miners can verify the work by retrieving only the relevant parts of the DAG and perform one hash operation.
- Current plan is to move to proof-of-stake soon (?)
    - The new protocol is called Casper.
    - The timeline has been pushed several times.
    - This is part of an upgrade known as Ethereum 2.0

Consensus

- Consensus is just like in Bitcoin, accept a block implicitly by mining on top of it.
- A new block is mined every 15 seconds on average.
  - The fast block generation rate means higher probability of having orphan blocks.
- The number of transactions in a block is specified by the block gas limit, i.e., the max total gas amount spent by all transactions in a block.
  - Currently this is limited to 8 million gas units.
- In case of forking, the longest chain is selected.

Uncle Blocks

- Miners are rewarded for mining as well as for including uncle blocks.
  - An uncle is an orphan block in Bitcoin; A valid block that was mined roughly at the same time with other valid blocks at the same height.
  - By including an uncle, the winning miner gets ⅛ of the mining rewards. The uncle miner gets ⅞ of the mining rewards. (currently a mining reward is 2 ETH, first uncle is 1.75 ETH).
  - A new block can reference up to 2 uncles.
- Why to include uncles?
  - Increase network security, more work is needed to remine.
    - Mainly to mitigate the consequences of fast block generation rate.
  - Reward smaller mining pools and individual miners for the work.
- Do not contribute in updating the state of the system.

Ethereum’s Blockchain I

- In its yellow paper, Ethereum's blockchain is defined as “cryptographically secure transactional singleton machine with shared-state.”.
    - The blockchain operates as a single machine responsible of tracking all transactions, i.e., a single truth about the system’s state.
    - The system state, or blockchain content, is shared across several machines or miners.
- Ethereum’s state machine changes state based on the transactions processed so far.
    - A state machine is a machine that reads an input and changes to a new state based on the output according to some transition function.

Ethereum’s Blockchain II

- This state machine starts with the genesis state (aka genesis block).
- Similar to Bitcoin, transactions are grouped into blocks, and these blocks are chained using their hash.
- Based on the transactions included in a block, the newly ined block defines a new state for the system.
    - An account state can contain the account balance, reputation, contract code associated with the account, or any digital information about the system.
    - A state it is a mapping between addresses and account states.
- In addition, a block contains an identifier for the new system state.
    - This ID is simply the root of the Merkle tree over all mappings in the state.

Ethereum's Blockchain III

- The mapping between addresses and accounts is stored in a state tree called Patricia Tree.
    - A combination of radix trees (or prefix trees/tries) and Merkle trees.
        - (For more information see: https://eth.wiki/en/fundamentals/patricia-tree).
    - The hash of the root node of the tree is stored in a block's header to reflect the new state of the system.
        - The full tree is stored off-chain.
- Each block header also contains hashes of the root nodes of the transactions tree, contract storage tree, and transaction receipt tree for all transactions included in the block.

Miner’s Rewards

- Similar to Bitcoin, miners have two sources of income:
  - Mining rewards (newly minted currency in each newly mined block).
    - Decrease over time to reduce inflation.
    - Block and uncle block miners collect fees.
  - Transactions fees in the form of computation cost in gas units.
    - Uncle miners do not collect such fees.