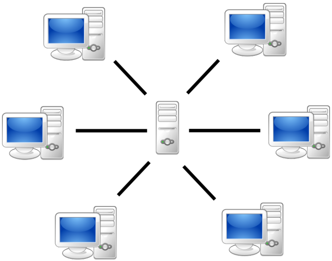
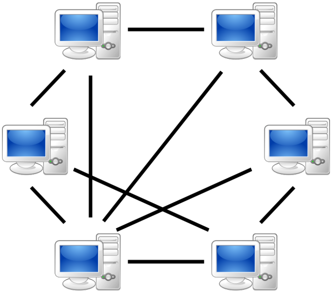
Improved scaling of the disk space taken by the Bitcoin blockchain

Benjamin Loison

Introduction to blockchains

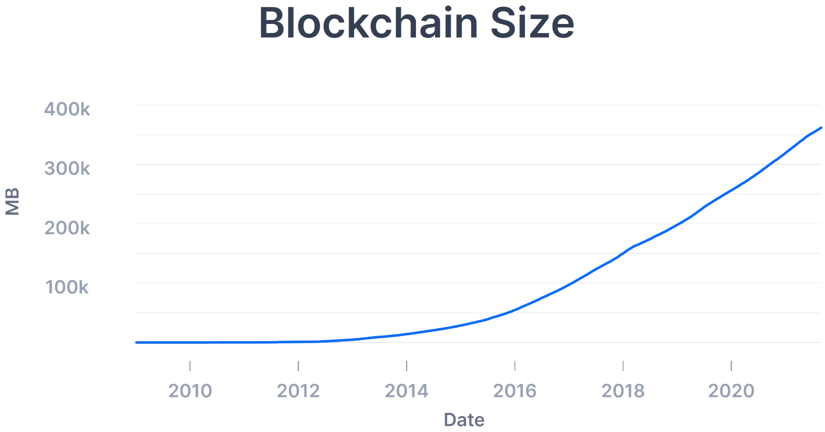


(a) Master-slave network



(b) Peer-to-peer network

The scalability problem



The idea of the internship

- Mining in Logarithmic Space 2021 Aggelos Kiayias, Nikos Leonardos and Dionysis Zindros

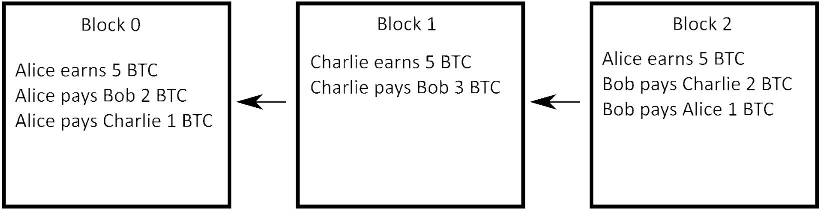
Transactions history

Person 1 pays	person 2	n Bitcoins
Network	Alice	5
Alice	Bob	2
Alice	Charlie	1
Network	Charlie	5
Charlie	Bob	3
Bob	Charlie	2
Network	Alice	5
Bob	Alice	1

Current state

Person has	n Bitcoins
Alice	8
Bob	2
Charlie	5

How Bitcoin works



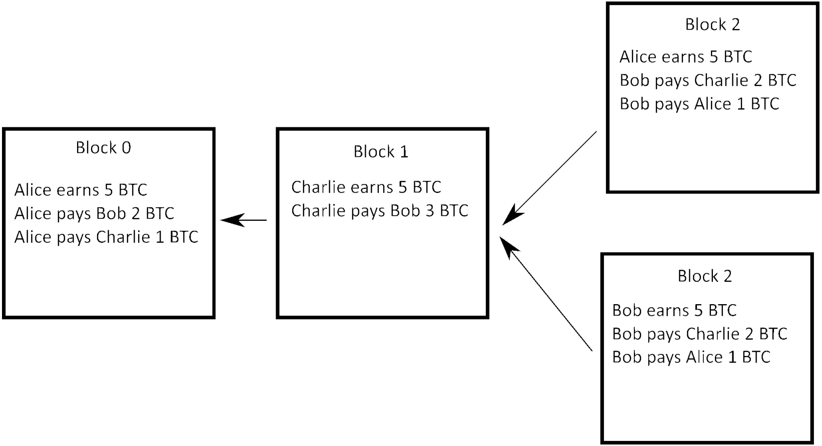
Mining blocks

Data:

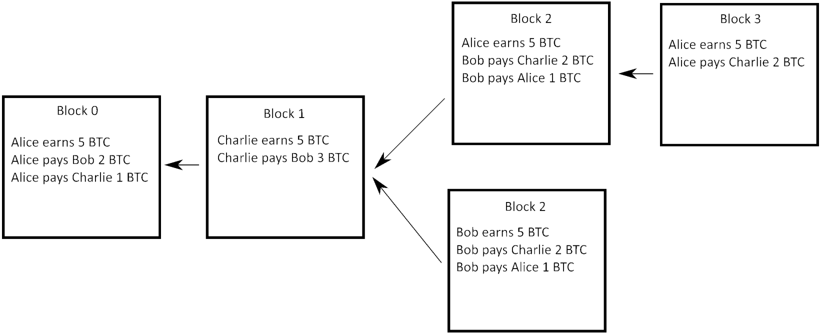
Block 0 n
Alice earns 5 BTC
Alice pays Bob 2 BTC
Alice pays Charlie 1 BTC

n	SHA-256² hash
0	6c7c2450bd52e950a3db47d8dc91cbdb04a792561759. . .
1	6442a403b0cd2bac7b3af363a342769d1955f9851d65. . .
. . .	. . .
86	00e9d707e8f386a73d2455cfa9c06d618285f03e434a. . .

The fork problem



The fork problem



The advantages of the theory

Proposal	Storage	Communication	Can mine?
BTC Full	$n(c+\delta)$	$n(c+\delta)$	yes
BTC SPV	nc	nc	no
Ethereum	$nc + k\delta + a$	$nc + k\delta + a$	yes
This work	$\mathit{poly}\log(n)c + k\delta + a$	$\mathit{poly}\log(n)c + k\delta + a$	yes

Table 1. A comparison of our results and previous work. n: the number of blocks in the chain; δ: size of transactions in a block; c: block header size; a: size of snapshot; k: common prefix parameter

Figure: Excerpt from the table on page 9 of "Mining in Logarithmic Space" (BTC means Bitcoin)

$n = 695590$, δ between 0 and 2 Mb, $c = 97$, $a = 4.24$ Gb, $k = 6$

The interlink set problem

Fig. 2. The interlinked blockchain. Each superblock is drawn taller according to its level. A new block links to all previous blocks that have not been overshadowed by higher levels in the meantime.

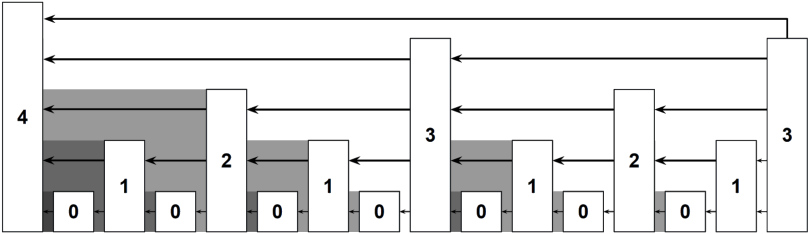


Figure: Set of "Mining in Logarithmic Space" pointers necessary for the proper execution of their approach

Some statistics

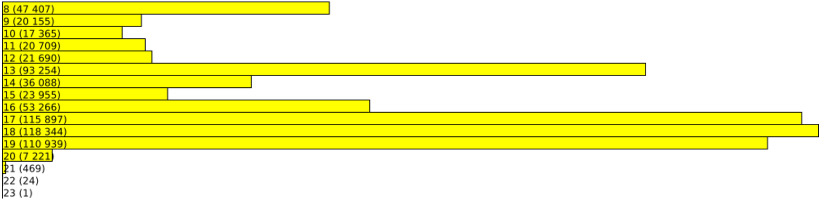


Figure: Distribution of Bitcoin block hashes by difficulty m (n) where m is the number of hexadecimal zeros at the beginning of the hash and n the number of hashes beginning precisely with m hexadecimal zeros

Some statistics

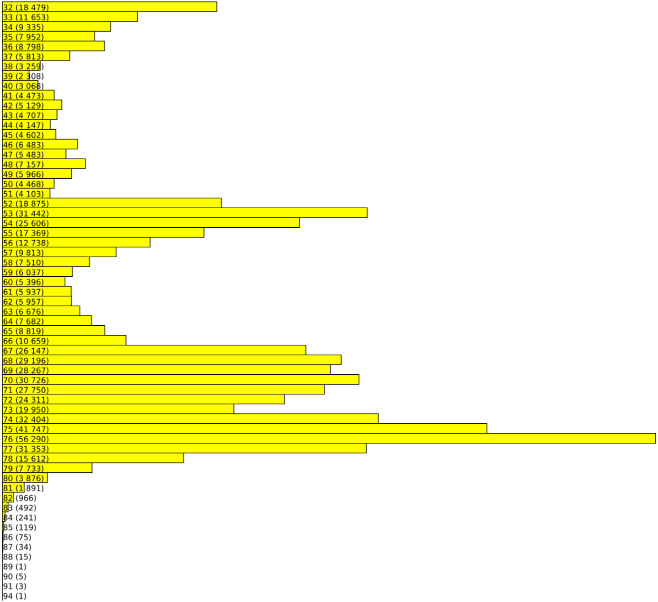


Figure: Distribution of Bitcoin block hashes by difficulty *m* (*n*) where *m* is the number of binary zeros at the beginning of the hash and *n* the number of hashes beginning precisely with *m* binary zeros

The compression algorithm

Algorithm 1 Chain compression algorithm for transitioning a full miner to a logspace miner. Given a full chain, it compresses it into logspace state.

function $\mathsf{Dissolve}_{m,k}(\mathcal{C})$
    $\mathcal{C}^* \leftarrow \mathcal{C}[: -k]$
    $\mathcal{D} \leftarrow \emptyset$
    if $|\mathcal{C}^*| \geq 2m$ **then**
        $\ell \leftarrow \max\{\mu : |\mathcal{C}^*{\uparrow^\mu}| \geq 2m\}$
        $\mathcal{D}[\ell] \leftarrow \mathcal{C}^*{\uparrow^\ell}$
        for $\mu \leftarrow \ell - 1$ down to 0 **do**
            $b \leftarrow \mathcal{C}^*{\uparrow^{\mu+1}}[-m]$
            $\mathcal{D}[\mu] \leftarrow \mathcal{C}^*{\uparrow^\mu}[-2m:] \cup \mathcal{C}^*{\uparrow^\mu}\{b:\}$
        end for
    else
        $\mathcal{D}[0] \leftarrow \mathcal{C}^*$
    end if
    $\chi \leftarrow \mathcal{C}[-k:]$
    return $(\mathcal{D}, \ell, \chi)$
end function
function $\mathsf{Compress}_{m,k}(\mathcal{C})$
    $(\mathcal{D}, \ell, \chi) \leftarrow \mathsf{Dissolve}_{m,k}(\mathcal{C})$
    $\pi \leftarrow \bigcup_{\mu=0}^{\ell} \mathcal{D}[\mu]$
    return $\pi\chi$
end function

Figure: Algorithm 1 of "Mining in Logarithmic Space" allowing to compress a blockchain.

C is the blockchain

$C^* \uparrow^\mu$ denotes blocks of exactly the same difficulty level μ of C^*

$C^* \uparrow^\mu \{b :\}$ denotes blocks of $C^* \uparrow^\mu$ newer than the block b

The results

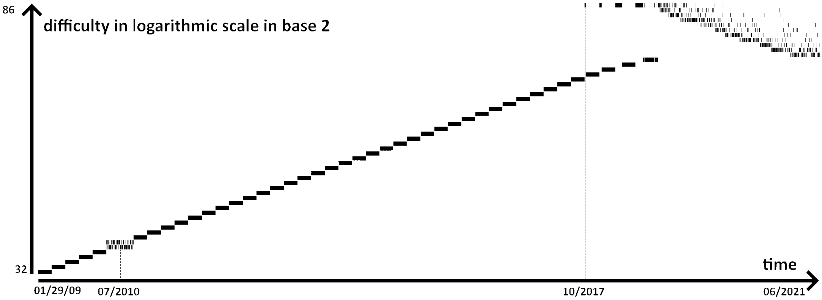


Figure: Distribution of the hashes of the blocks selected by the algorithm 1, where each block has a width of 1 pixel

Sources of illustrations

- Page 2: Wikipedia: peer-to-peer
- Page 3: Blockchain.com