Technologia blockchain

Henryk Budzisz

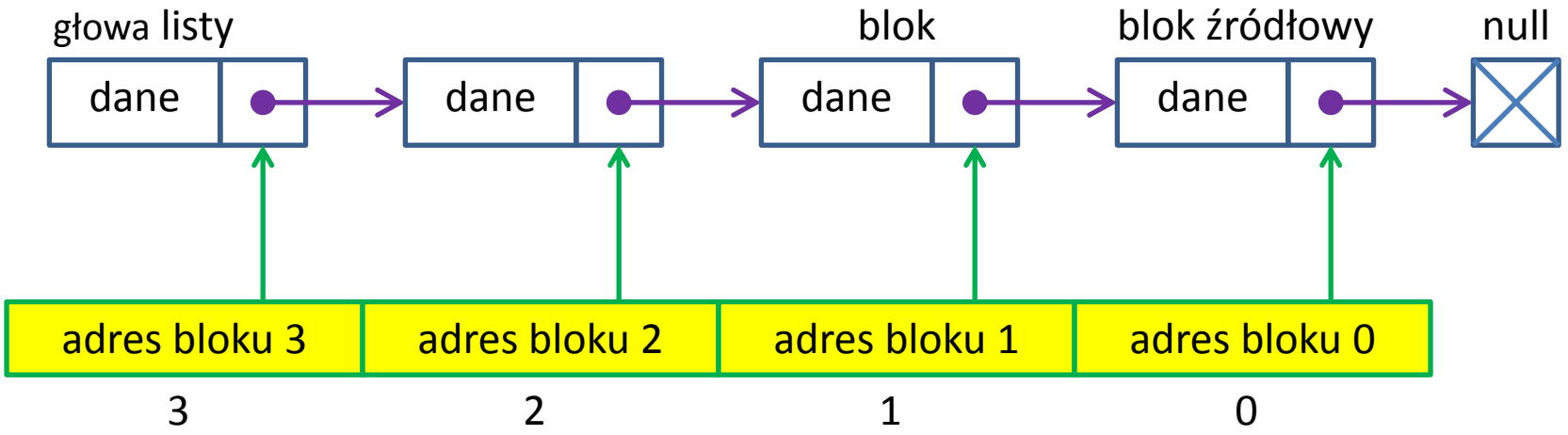
Koszalin 2018

Plan prezentacji

- struktura łańcucha
- transakcje
- portfel – cyfrowy podpis transakcji
- ochrona danych – zagadka kryptograficzna
- sieć „peer-to-peer”
- bitcoin
- co dalej?

Struktura danych

Z punktu widzenia struktury danych *blockchain* jest listą jednokierunkową.



Elementy tej listy nazywane są blokami, a link łączący je jest **kryptograficznym skrótem** danych.

Funkcja skrótu (hash)

SHA-2 (**Secure Hash Algorithm**) to zestaw kryptograficznych funkcji skrótu zaprojektowany przez National Security Agency (NSA) i opublikowany w 2001 roku.

Funkcje tworzą niepowtarzalny i nieodwracalny skrót wiadomości.

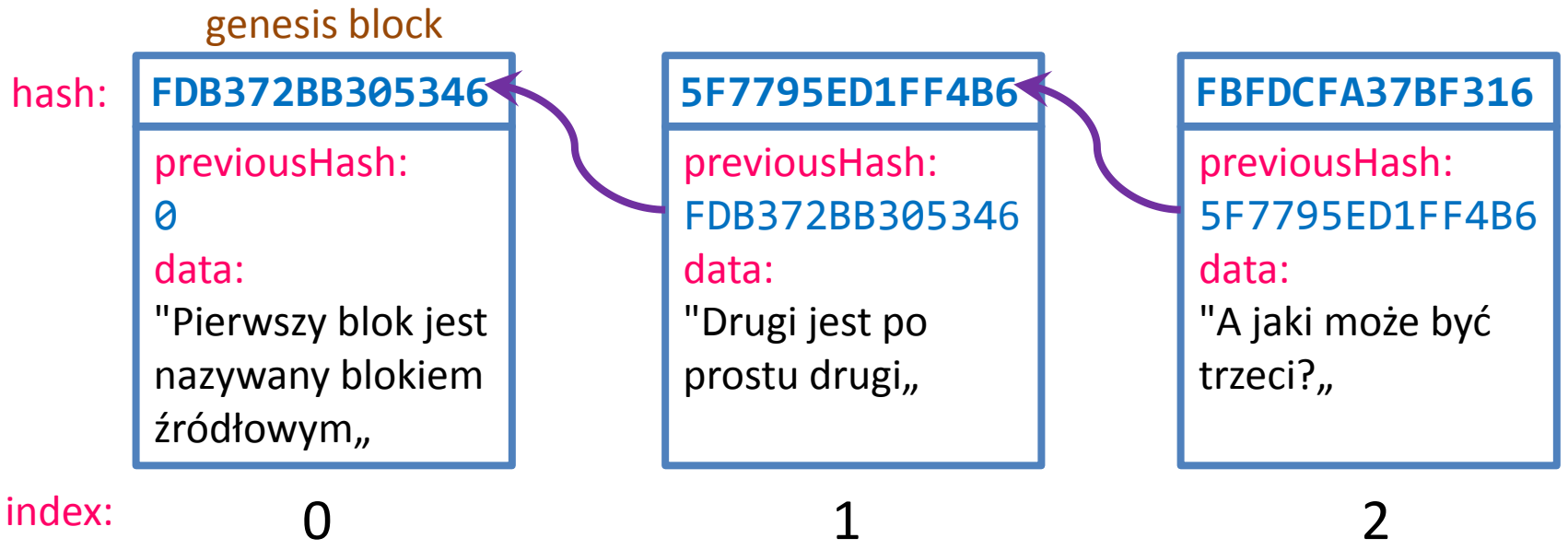
Wiadomość jest dzielona na pakiety o stałym rozmiarze (512 lub 1024 bitów), które następnie są mieszane. W procesie mieszania (64 lub 80 rund) wykorzystywane są funkcje: +, *and, or, xor, shr, rot*. Eksperymentalna wydajność przekracza 100 MB/s.

Skrót ma stały rozmiar zależny od wersji. Powszechnie stosowana wersja SHA-256 daje skrót 256 bitowy (64 znaki w zapisie heksadecymalnym).

Funkcje skrótu zostały zaimplementowane w wielu bibliotekach programistycznych.

Przykład: SHA256test

Struktura prostego łańcucha bloków



Zmiana zawartości jakiegoś bloku wymusza **zmianę skrótu do tego i <u>wszystkich następnych bloków</u>**. Jeżeli do tego nie dopuścimy, dane pozostaną nienaruszone.

Przykład: SimpleBlockchain.

Zastosowania łańcucha bloków

Najbardziej powszechnym obszarem zastosowań łańcucha bloków jest ustalanie własności i zarządzanie nią, ale nie tylko.

W tym kontekście *blockchain* pełni rolę rejestru danych dotyczących przenoszenia własności. Dane zapisane w rejestrze nie mogą ulec zmianie. Funkcjonuje jako struktura typu „tylko do dopisywania”.

Historia przenoszenia własności pewnego dobra umożliwia ustalenie jego bieżącego właściciela. Tylko uprawniony właściciel może przenieść własność dobra na inną osobę.

W obecnych zastosowaniach tym dobrem jest najczęściej kryptowaluta (bitcoin, ethereum, ripple, litecoin, monero itd.)

Dokumentowanie własności

Dwa konkurencyjne sposoby:

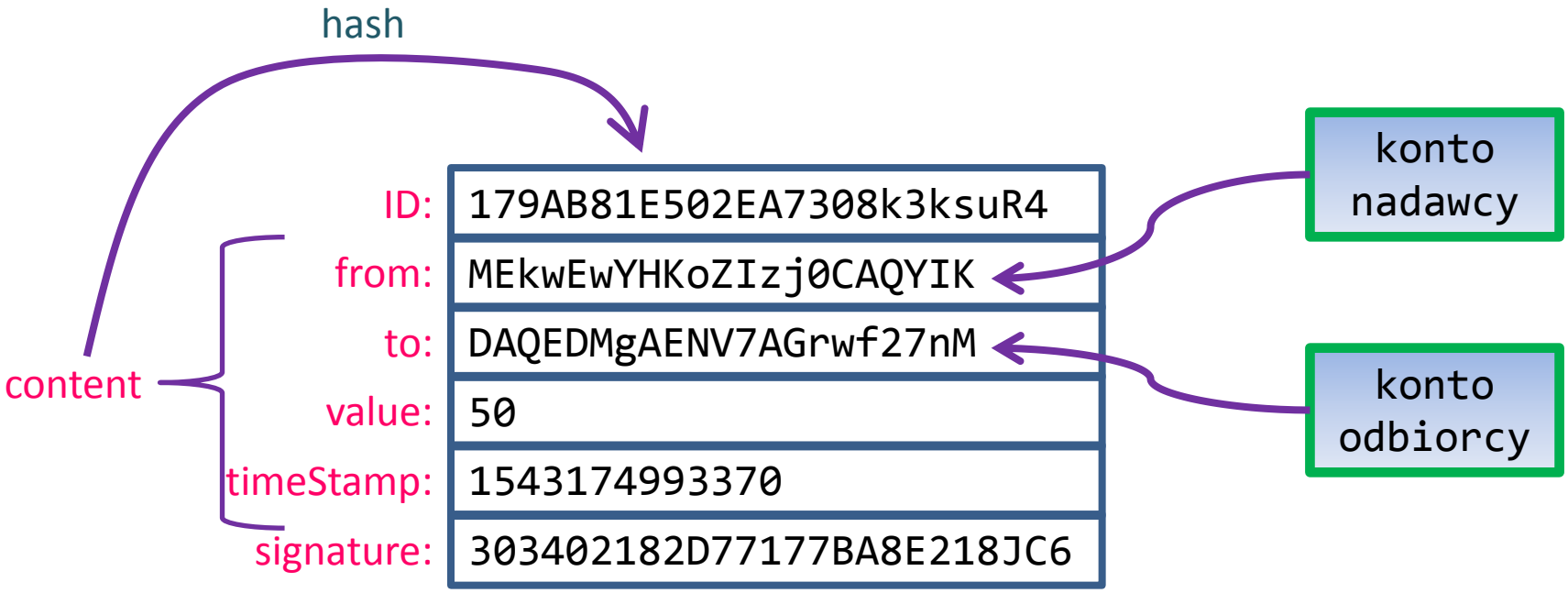
- poprzez **dane inwentaryzacyjne**, opisujące obecny stan własności.
- poprzez **dane transakcyjne** opisujące przeniesienia własności.

Cała historia danych transakcyjnych przechowywana w rejestrze staje się ścieżką audytu, która dostarcza dowodów na to, w jaki sposób określone osoby weszły w posiadanie określonych dóbr.

Dane inwentaryzacyjne można wytworzyć agregując dane transakcji.

Blockchain jest uporządkowanym czasowo rejestrem transakcji

Transakcja



Transakcja jest opisem przeniesienia własności.

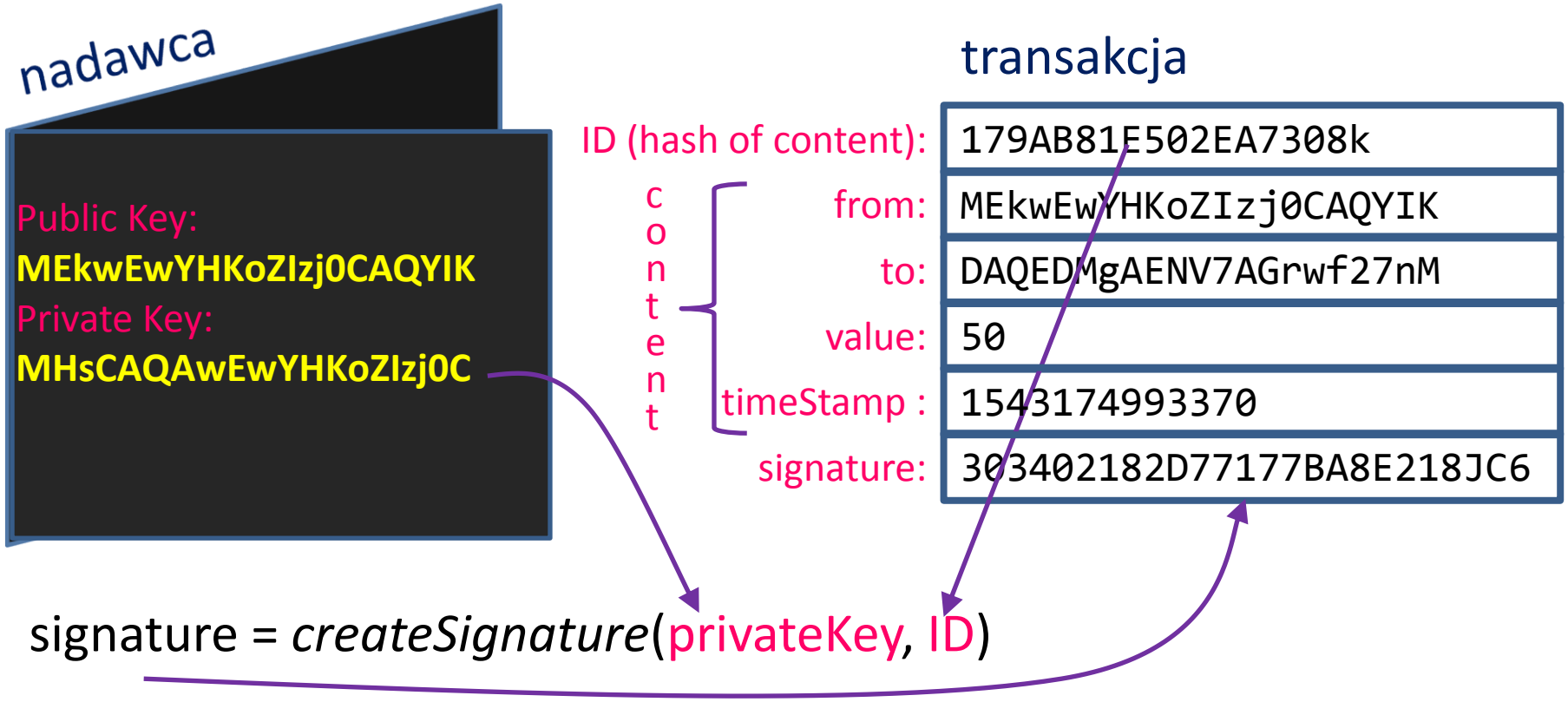
Konto użytkownika (portfel)

Klucz prywatny służy do podpisania transakcji. Znany jest tylko właścicielowi konta.

Klucz publiczny umożliwia weryfikację podpisu. Jest równocześnie adresem konta (portfela), widocznym w transakcji.

Dwa podpisy tej samej wiadomości mogą być identyczne lub różne – zależnie od stosowanego algorytmu.

Podpis cyfrowy



Podpis może sprawdzić każdy do kogo dotarła transakcja – przy użyciu klucza publicznego nadawcy (from)

boolean *verifySignature*(from, signature, hash(content))

Portfel jest pusty

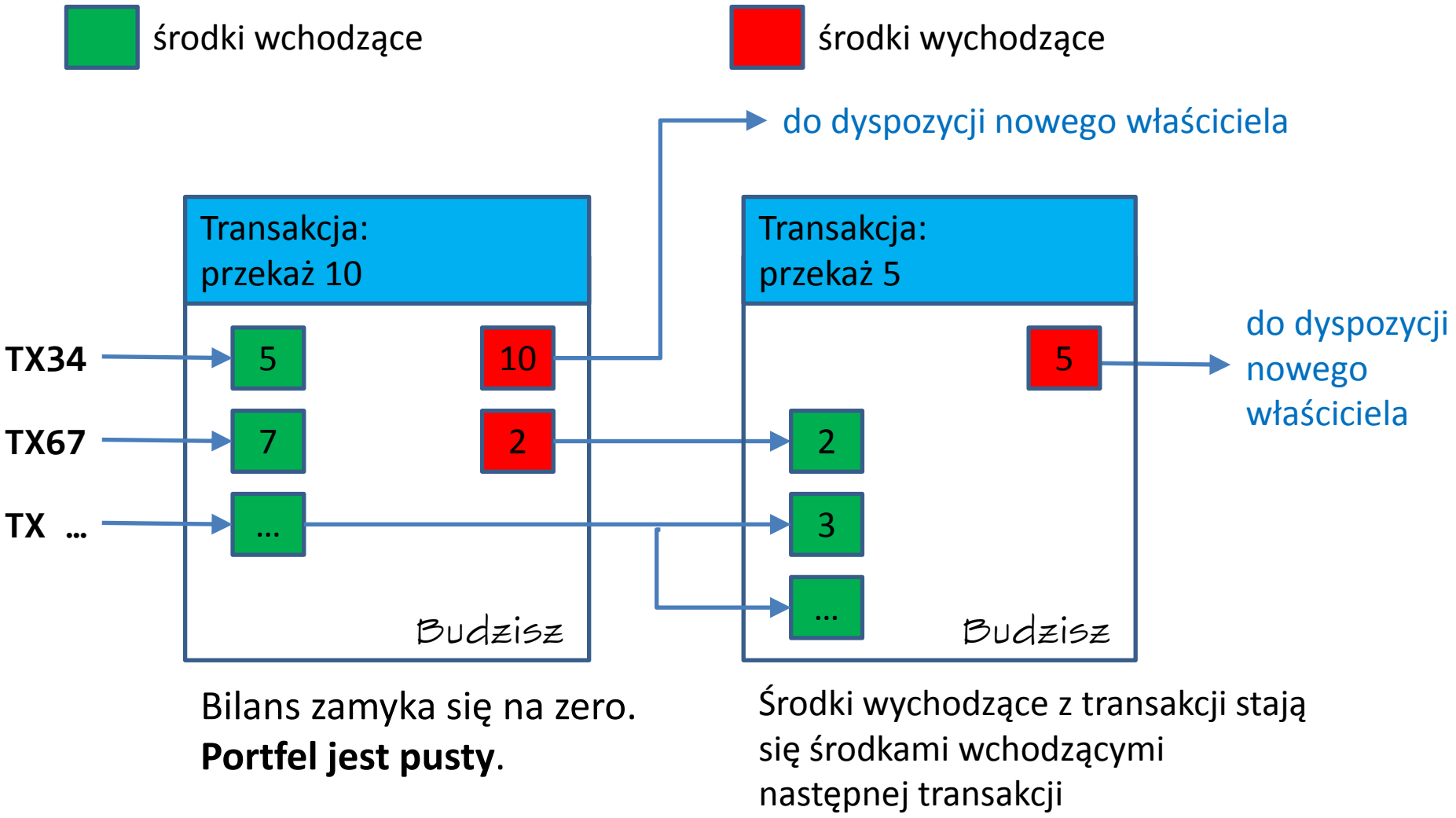
Technologia blockchain definiuje pojęcie własności od nowa. Właściciel konta (portfela) ma dostęp tylko do transferu pewnej własności na swoje konto.

Może tą własność zatrzymać, przechowując identyfikator transakcji, albo też przekazać na inne konto (swoje lub czyjeś).

Portfel służy tylko do zorganizowania procesu przenoszenia własności. Do wyrażenia zgody właściciela na przekazanie środków poprzez udostępnienie podpisu cyfrowego

Żeby przekazać coś komuś, trzeba to coś najpierw otrzymać.

Łączenie i dzielenie środków



Bilans zamyka się na zero.
Portfel jest pusty.

Środki wychodzące z transakcji stają się środkami wchodzącymi następnej transakcji

Patrz: projekt *KoszalekChain*, *TransactionOutput*, *TransactionInput*, *Transaction*

Jakimi środkami dysponuję?

Właściciel konta jest anonimowy. Znany jest tylko adres konta (klucz publiczny), np.: MEkwEwYHKoZIzj0CAQYIKoZIzj0DAQEDMgAE8SjZrprhk9IGiB4

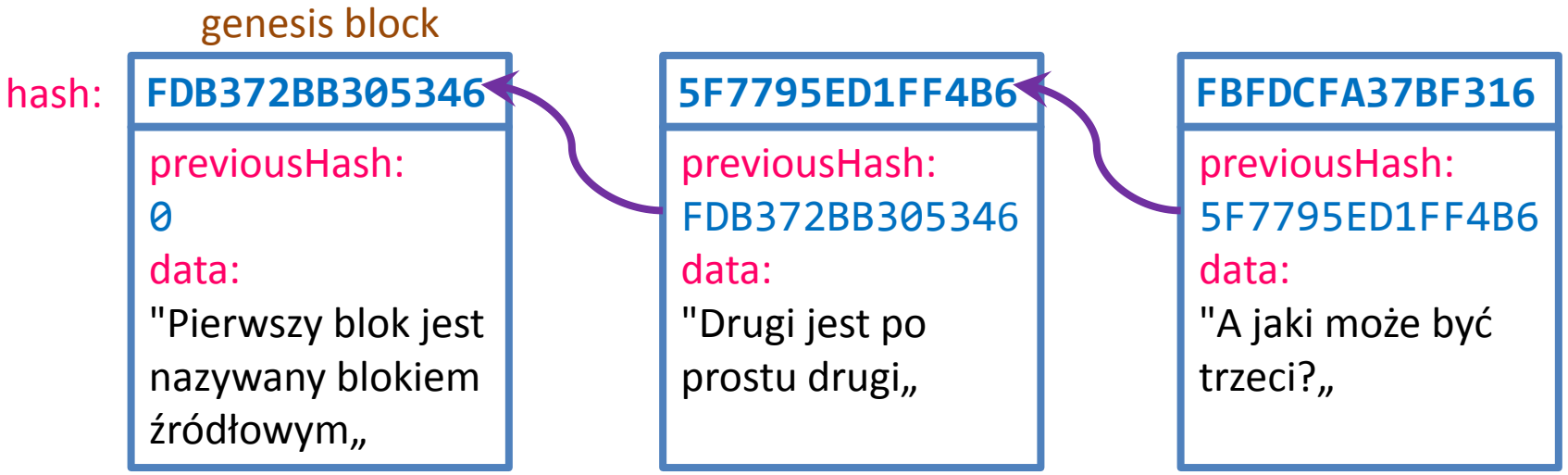
Nie ma żadnych możliwości powiadomienie właściciela o transakcji dotyczącej jego konta.

Przed wykonaniem transakcji sam musi sprawdzić czy istnieją transakcje zasilające jego konto, posługując się listą UTXO’s (Unspent Transaction Outputs).

Suma środków na tej liści daje środki, które aktualnie są dostępne dla wszystkich.

Przechowywanie transakcji

Przypomnienie struktury łańcucha bloków

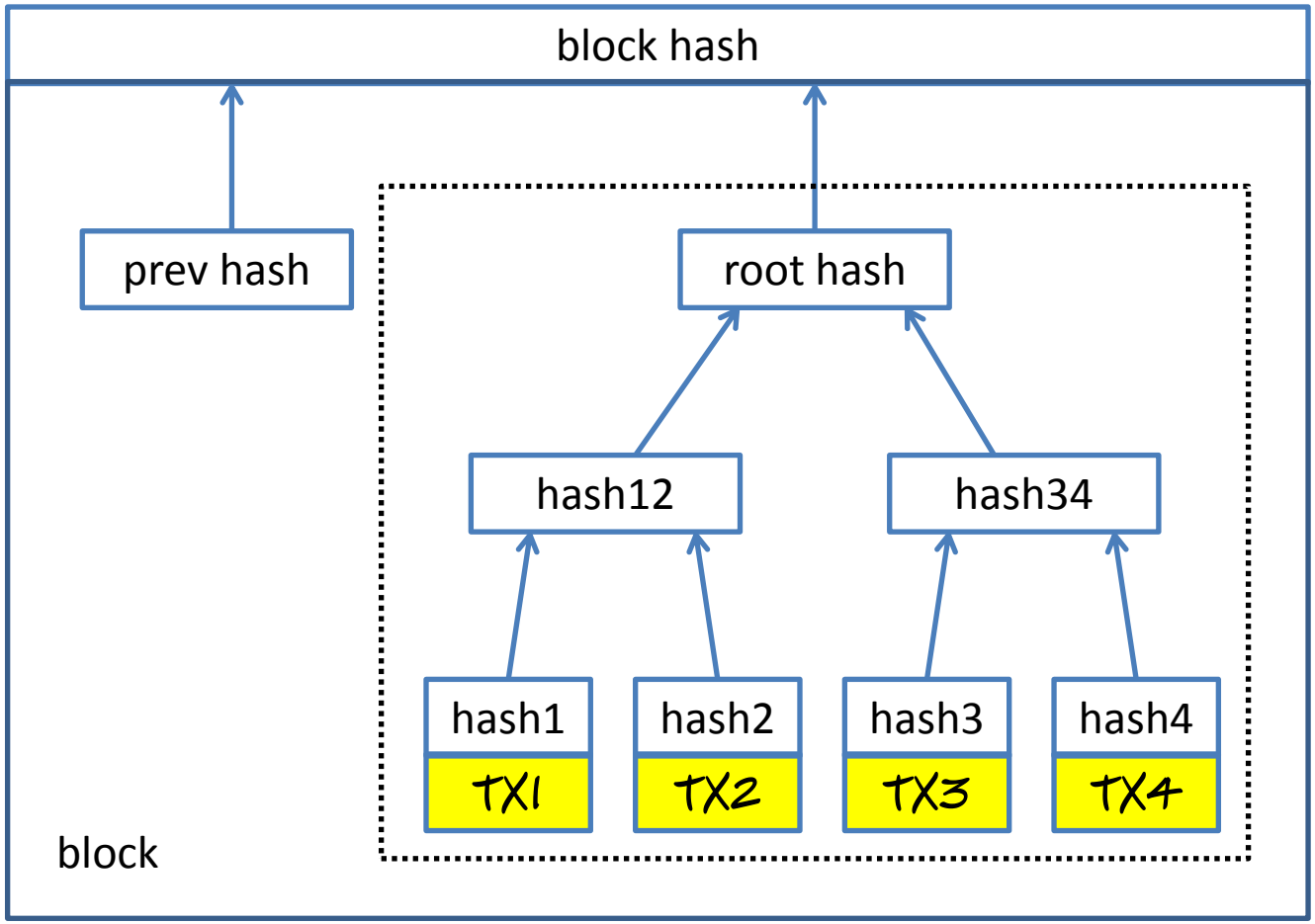


Przypomnienie: zmiana zawartości jakiegoś bloku wymusza **zmianę skrótu do tego i <u>wszystkich następnych bloków.</u>**

Tekst zastąpimy listą transakcji

Drzewo Merkle'a

binarne drzewo skrótów



W bloku przechowuje się tylko korzeń drzewa.

n	liczba TX
2	4
8	256
10	1024
12	4096

Patrz: projekt *KoszałekChain*, *Util.getMerkleTree*

Dodawanie nowego bloku do łańcucha

- dodać transakcje do bloku – liczba transakcji określa wielkość bloku 👁
- zbudować drzewo Merkle'a i zachować jego korzeń (*merkleRoot*) 👁
- uzyskać skrót do ostatniego bloku w łańcuchu (*previousHash*) 👁
- wyznaczyć stempel czasu (*timeStamp*) 👁
- obliczyć skrót z nagłówka bloku (hash) 👁
  nagłówek = previousHash + timeStamp + merkleRoot
- dodać blok do listy bloków tworzących *blockchain* 👁

Patrz: projekt *KoszałekChain* klasa *Block* i funkcja *addBlock* w klasie *Chain*

Ochrona danych

Cel

Cała historia danych transakcyjnych przechowywanych przez łańcuch bloków musi pozostać niezmienna, aby była godnym zaufania źródłem w zakresie ustalania kwestii związanych z własnością.

Problem

W środowisku sieciowym istnieje ryzyko, że nieuczciwi uczestnicy systemu mogliby manipulować historią danych transakcyjnych lub fałszować ją dla własnych korzyści.

Rozwiązanie

Aby zapewnić niezmienność historii transakcji, należy doprowadzić do tego, aby wprowadzanie zmian w historii wiązało się z zaporowymi kosztami, odwodzącymi wszystkich od podejmowania takich działań.

Skutki ataku

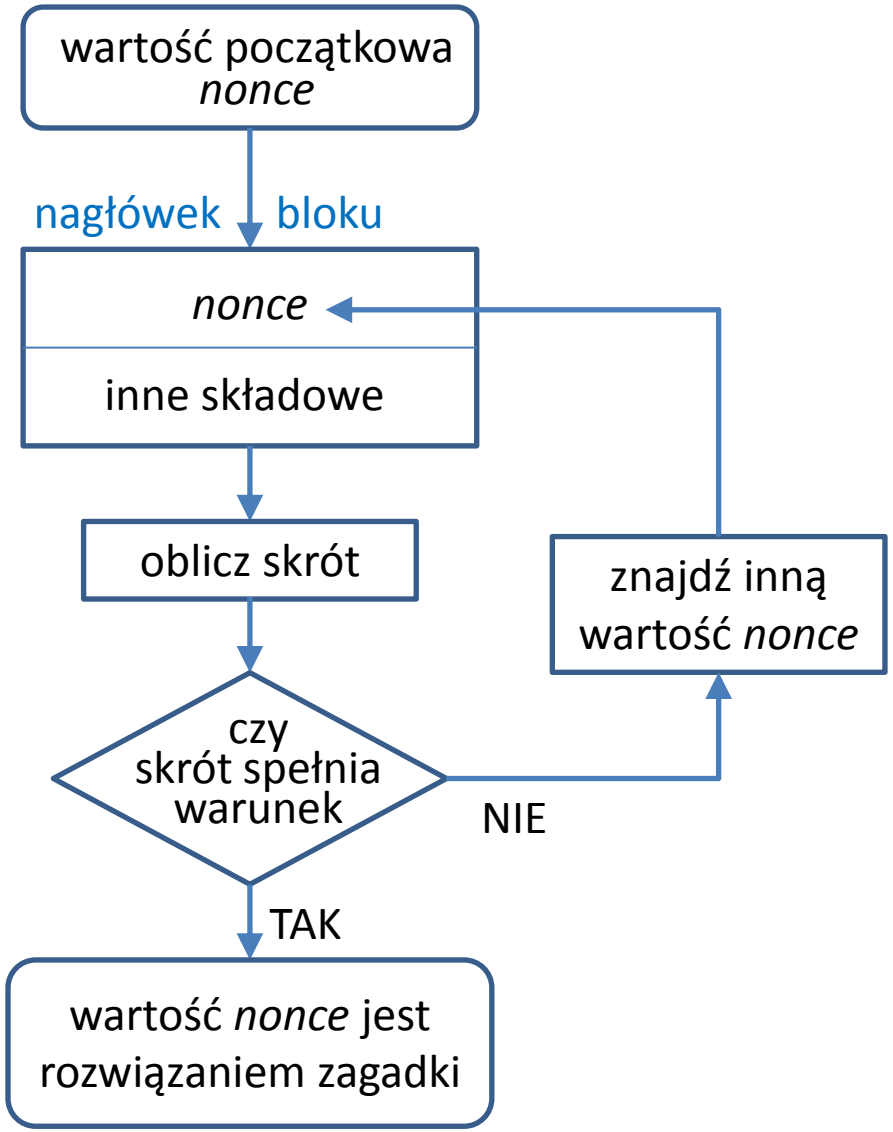
Zmiana zawartości transakcji – tak aby była niewykryta przez innych użytkowników – wymaga:

- zmianę skrótu tej transakcji
- zmianę podpisu cyfrowego tej transakcji
- zmianę korzenia drzewa Merkle'a, w którym transakcja się znajduje
- zmianę skrótu bloku, bo zmienił się korzeń
- zmianę skrótu następnego bloku, bo zmienił się skrót poprzedniego
- itd. aż do najnowszego bloku

Struktura danych łańcucha bloków wymusza przepisywanie historii, aby móc wprowadzić zmiany. Jest to dosyć złożone, ale wykonalne.

Zagadka kryptograficzna

Znajdź liczbę (nonce), która połączona z pozostałymi składowymi nagłówka bloku da skrót spełniający pewien warunek



Warunek dla skrótu

Skrót SHA-256, na który nie nakłada się żadnych ograniczeń może przyjąć jedną z $2^{256}-1 \approx 10^{77}$ wartości.
Najprostszym sposobem zawężenia tego zbioru wartości jest sformułowanie relacji:

$$\text{skrót} < \text{max} < 16^{64}$$

Ponieważ skrót przedstawiany jest w zapisie hexadecymalnym, wartość max wyznacza się przez operację na łańcuchu znaków, wstawiając zera na pierwszych n pozycjach oraz F na pozostałych, wówczas $\text{max}=16^{64-n}-1$.
Przykład dla n=5, max=00000FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF...

Wówczas prawdopodobieństwo znalezienia skrótu spełniającego warunek wynosi 16^{-n}, a wartość n nazywana jest poziomem trudności.

Zamiast wyznaczać wartość relacji, wystarczy sprawdzić czy obliczony skrót ma w zapisie szesnastkowym zera na pierwszych n pozycjach.

n	16^{-n}	max
0	1	10^{77}
1	0,0625	$7*10^{75}$
2	0,0039	$4,5*10^{74}$
10	$9*10^{-13}$	10^{65}
20	$8*10^{-25}$	10^{53}

Algorytm rozwiązywania zagadki

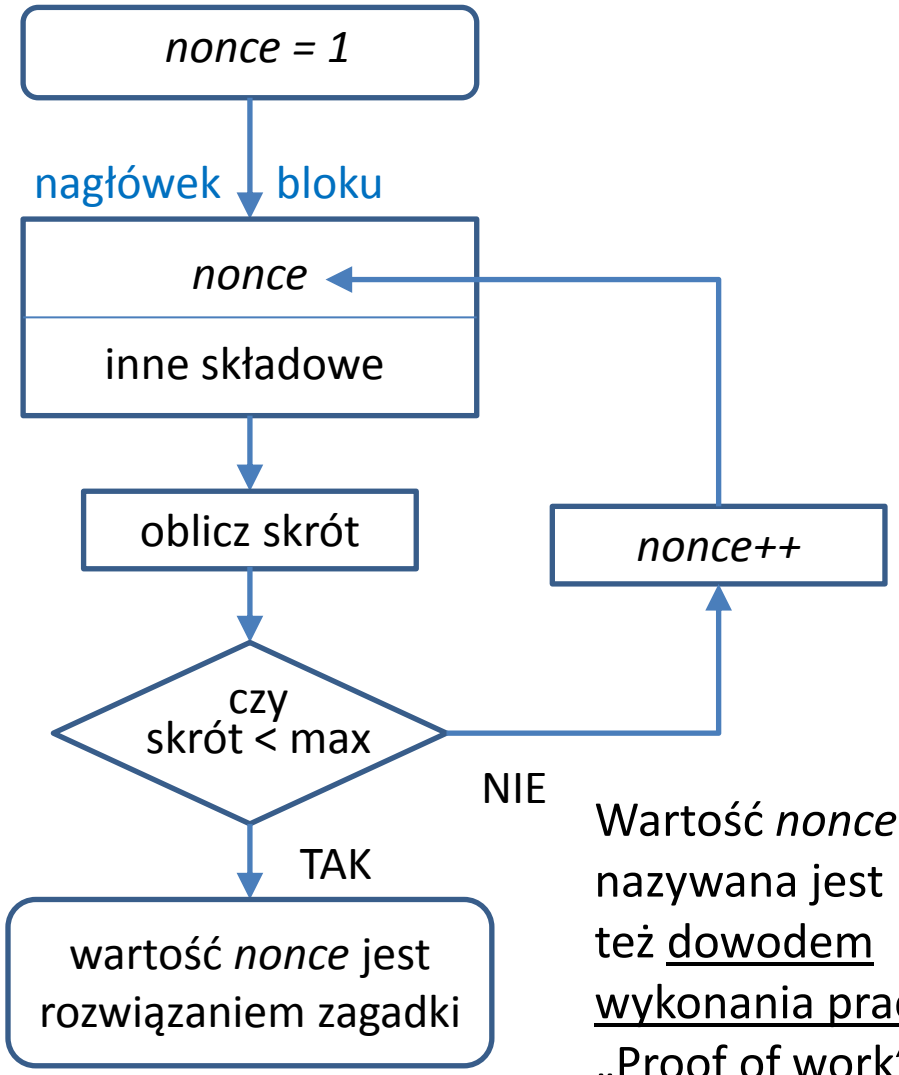
Nie ma żadnego sposobu żeby przewidzieć jaka powinna być wartość *nonce*, wobec tego można zastosować prostą iterację.

Jest wiele wartości *nonce*, dla których skrót spełnia relację. Wystarczy znaleźć jedną.

Rozwiązanie (nonce) i poziom trudności (n) umieszcza się w nagłówku bloku – są jawne.

Sprawdzenie czy zagadka została rozwiązana wymaga jednorazowego obliczenia skrótu.

Patrz: projekt *Mining*.



Wartość *nonce* nazywana jest też dowodem wykonania pracy „Proof of work” (PoW).

Wyścig

Wyścig pomiędzy łańcuchem uczciwym i łańcuchem atakującego można opisać jako dwumianowy proces błądzenia losowego.

Postęp atakującego będzie miał rozkład Poissona o następującej oczekiwanej wartości:

$$\lambda = z(q/p), \quad \text{gdzie:}$$

p - prawdopodobieństwo, że uczciwy węzeł znajdzie kolejny blok,

q - prawdopodobieństwo, że napastnik znajdzie następny blok,

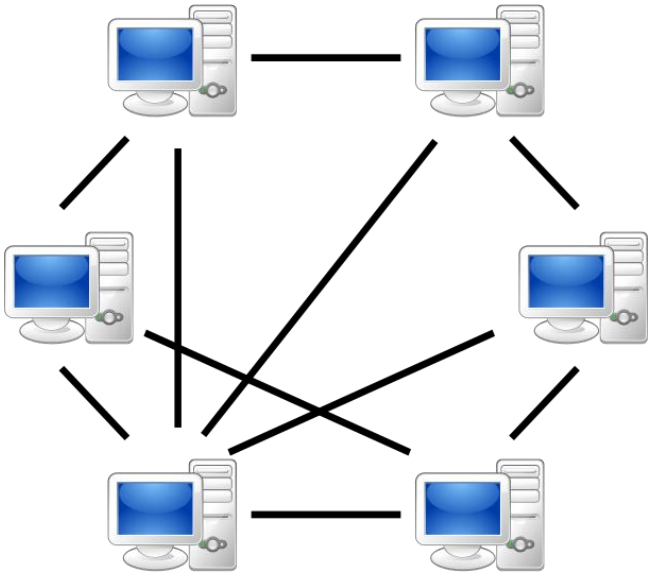
z – liczba bloków do nadrobienia.

Atakujący musiałby dysponować większą mocą od sumy mocy pozostałych węzłów, aby wyścig kiedykolwiek wygrać (51% hashrate attack).

5 stycznia 2019 udał się atak 51% na kryptowalutę Ethereum Classic – straty wyniosły ok. 1 mln USD.

Sieć „peer-to-peer” (P2P)

P2P to model komunikacji w sieci komputerowej zapewniający węzłom te same uprawnienia.



Węzeł pełni jednocześnie rolę klienta i serwera. W roli serwera każdy węzeł udostępnia sieci swoją moc obliczeniową i przestrzeń dyskową.

W technologii blockchain każdy węzeł przechowuje kopię łańcucha, a węzły konkurują między sobą w procesie powiększania łańcucha. Zachętą jest nagroda.

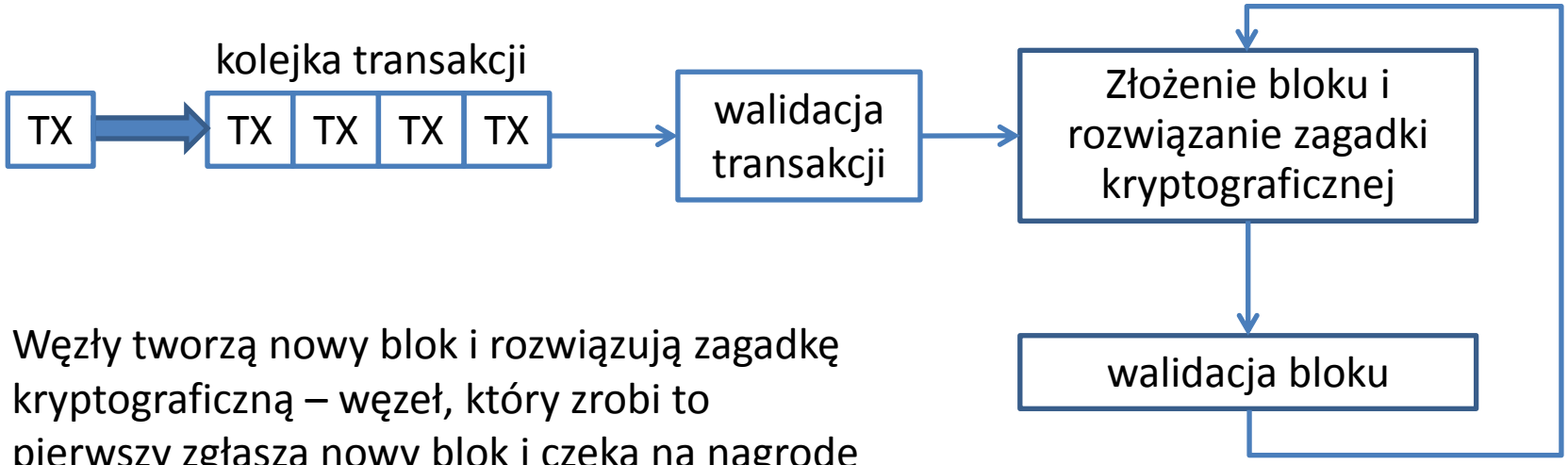
Jak to działa

Utrzymywanie istniejących połączeń
Każdy węzeł w sieci przechowuje listę równorzędnych węzłów, z którymi się komunikuje. Lista zawiera tylko podzbiór wszystkich węzłów składających się na cały system. Każdy węzeł na bieżąco weryfikuje, czy równorzędne węzły są ciągle dostępne – jeżeli nie są – usuwa je z listy.

Nawiązywanie nowych połączeń
Każdy węzeł może zażądać dołączenia go do systemu, wysyłając odpowiedni komunikat do dowolnego węzła. Węzeł, do którego skierowane zostało żądanie, dodaje jego adres do własnej listy węzłów i w odpowiedzi wysyła potwierdzenie. Po otrzymaniu odpowiedzi pierwszy węzeł także dodaje adres nadawcy do własnej listy węzłów. Powstaje nowe połączenie.

Rozpowszechnianie nowych informacji
Nowe informacje (np. nowe dane transakcyjne i nowe bloki) są dystrybuowane do węzłów na liście, w miarę pojawiania się. Każdy węzeł podłączony do systemu w końcu otrzymuje wszystkie wiadomości. Duplikaty są na bieżąco usuwane.

Walidacja danych



Węzły tworzą nowy blok i rozwiązują zagadkę kryptograficzną – węzeł, który zrobi to pierwszy zgłasza nowy blok i czeka na nagrodę

Wszystkie pozostałe węzły przerywają pracę nad budowaniem nowego bloku (bo już przegrały konkurs szybkości) i przystępują do walidacji zgłoszonego bloku.

W określonym momencie każdy z węzłów systemu realizuj jedno z dwóch zadań:

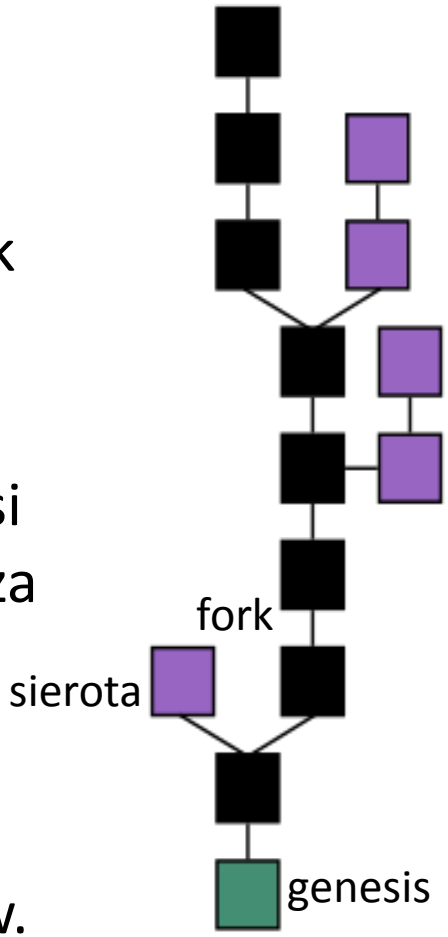
1. Ocenia nowy blok, który został utworzony i przesłany przez jednego z uczestników systemu.
2. Stara się zostać następnym węzłem, który stworzy nowy blok i prześle go do oceny wszystkim pozostałym.

Rozgałęzianie się drzewa

Opóźnienia w nadsyłaniu nowych bloków przez sieć oraz sytuacje, w których dwa węzły tworzą nowe bloki niemal równocześnie, powodują, że część węzłów otrzymuje jako pierwszy jeden blok i przystępuje do weryfikacji, a część węzłów inny blok.

Struktura łańcucha rozdwaja się - zjawisko to nosi nazwę *fork.* Wygrywa ta gałąź, która jako pierwsza urośnie.

Węzły sieci pracujące nad gałęzią krótszą porzucają ją, bo nie mają szans na przegonienie gałęzi dłuższej. Bloki tworzące gałąź krótszą – tzw. sieroty – są odrzucane, a nagroda odbierana.



Bitcoin

Koncepcja bitcoina przedstawiona została w tzw. manifeście opublikowanym w 2008 roku przez Satoshi Nakamoto. W 2009 roku opublikowane zostało oprogramowanie obsługujące blockchain z transakcjami w tej kryptowalucie.

System pozwala na anonimowe posiadanie własności oraz jej transfery.

Początkowa nagroda za dodanie bloku wynosiła 50 BTC. Co 4 lata zmniejsza się o połowę, aby zachować wartość kryptowaluty. Obecnie wynosi 12,5 BTC. Poziom trudności jest dobierany automatycznie – tak aby budowa bloku trwała ok. 10 min. Całkowita liczba bitmonet dąży asymptotycznie do granicy 21 milionów, co nastąpi ok. 2040 roku.

W wielu państwach istnieje możliwość płatności za usługi lub dobra tą walutą, nawet przy zakupie drobnych rzeczy.

22 maja – dzień Bitcoinowej Pizzy

22 maja 2010 roku pomiędzy dwoma użytkownikami forum dyskusyjnego *BitcoinTalk*.org doszło do pierwszej transakcji z użyciem BTC.

Przedmiotem transakcji były dwie pizze z papryką, cebulą, pieczarkami, pomidorami i pepperoni. Wartość transakcji to 10 000 BTC (170 mln zł po dzisiejszym kursie). Ale mają wspomnienia …

Dane liczbowe: grudzień 2018

- liczba węzłów – ok. 11 000
- liczba portfeli – ok. 22mln
- liczba BTC wydobytych – ok. 17 400 000
- ok. 9 600 000 BTC (55%) znajduje się w portfelach, których saldo przekracza 200 BTC (1,25 mln $)
- liczba bloków – ponad 550 000
- objętość łańcucha – ponad 100 GB
- ocenia się, że nawet 30% BTC zostało utraconych lub bezpowrotnie zgubionych.

Przegląd bloków: https://www.blockchain.com/pl/explorer

Świat przestępczy

Hello!

I'm is very good programmer, known in darkweb as frederik22.
I hacked this mailbox more than six months ago, through it I infected your operating system with a virus (trojan) created by me and have been spying for you a very long time.

I have access to all your accounts, social networks, email, browsing history.
Accordingly, I have the data of all your contacts, files from your computer, photos and videos.

During your pastime and entertainment there, I took screenshot through the camera of your device, synchronizing with what you are watching.
Oh my god! You are so funny and excited!

I think that you do not want all your contacts to get these files, right?
If you are of the same opinion, then I think that $248 is quite a fair price to destroy the dirt I created.

Send the above amount on my BTC wallet (bitcoin): 1JRCbCH9E3iLhSXPTqtkgfAsJNT2xD74C5
As soon as the above amount is received, I guarantee that the data will be deleted, I do not need it.

Czy kopanie kryptowalut może być biznesem?

Odpowiedź na pytanie zawarta jest w koncepcji łańcucha bloków. Musi istnieć nagroda za rozwiązanie zagadki kryptograficznej, bo nie byłoby chętnych do tworzenia nowych bloków.

Ten wygrywa (dostaje nagrodę),
kto dysponuje większą mocą obliczeniową.

Koparki kryptowalut

Największą firmą tworzącą sprzęt do wydobywania kryptowalut stanowi aktualnie **Bitmain** (Chiny).

Sprzęt Bitmainu oparty jest o układy scalone ASIC. Jest wyspecjalizowanym sprzętem zaprojektowanym do kopania kryptowalut w oparciu o konkretny algorytm mieszający (hashing algorithm), taki jak SHA-256 lub Scrypt (na których odpowiednio działają Bitcoin i Litecoin).

Wydajność koparek mierzy się w haszach na sekundę (H/s). Wprawny rachmistrz ma wydajność 0,67 H/dobę.

Antminer S15

Cena: 1475$
Wydajność: 28 TH/s
Pobór mocy: 1600W
Zużycie roczne: 14MWh (8400 zł)

Kopalnie kryptowalut

Aktualnie kopanie bitcoinów wymaga sprzętu o gigantycznej mocy obliczeniowej. Wewnątrz sieci, w celu zwiększenia wydajności wydobycia, górnicy łączą siły tworząc **kopalnie**.

Koparki kryptowalut na świecie zużywają 140 TWh/rok – prawie tyle co cała Polska (164 TWh/rok).

Ograniczenia łańcucha bloków

- **brak prywatności** – ujawnienie konta ujawnia wszystkie transakcje przeprowadzone na tym koncie,
- **model bezpieczeństwa** – jedynym zabezpieczeniem portfela jest klucz prywatny; jego utrata, to katastrofa,
- **wysokie koszty** – cena bezpieczeństwa historii transakcji to konieczność rozwiązania zagadki kryptograficznej
- **ograniczona skalowalność** – mining spowalnia szybkość przetwarzania danych ograniczając skalowalność,
- **ukryte scentralizowanie** – kopalnie eliminują pozostałe węzły – powstaje oligopol i pokusa nadużyć,
- **brak elastyczności** – łańcuch bloków został zaprojektowany do działania bez nadzoru, a więc też bez możliwości wprowadzania zmian technologicznych,
- **wielkość krytyczna** – w małych sieciach intruz z dużą mocą obliczeniową może przeprowadzić skuteczny atak.

Co dalej?

Źródłem ograniczeń są dwa podstawowe konflikty:

- **transparentność** kontra **prywatność** – związany z dostępem do odczytu danych
- **bezpieczeństwo** kontra **szybkość** – związany z możliwością zapisu danych

Przedstawiona koncepcja łańcucha bloków zapewnia transparentność i bezpieczeństwo, kosztem prywatności i szybkości.

Ograniczenie dostępu do odczytu danych i/lub możliwości zapisu danych prowadzi do nowych koncepcji – prywatnego łańcucha bloków.

Nowe koncepcje

- **Lighting Network** – umożliwia wykonywanie mikrotransakcji pomiędzy właścicielami wspólnego konta, z pominięciem łańcucha głównego, co radykalnie zwiększa skalowalność sieci.
- **Sharding** – podział sieci węzłów przy użyciu złożonego algorytmu na podsieci współpracujące ze sobą (od *shards=odłamki*)
- **Enigma** – rozwiązuje kwestię prywatności poprzez technologię o nazwie „tajne umowy” (*secrets contracts*). Transakcje są szyfrowane. Węzły mogą je przetwarzać, ale nie rozpoznają danych na których operują.
- **DAG** (Direct Acyclic Graph) – Skierowany Graf Acykliczny wprowadza inną koncepcję rozproszonego rejestru, która już nie jest oparta na łańcuchu bloków.

Bez komentarza