8. Blockchain

- Einführung
- Elemente der Blockchain
  - Blockkette / Ledger
  - Double Spending Anomalie
- Konsensus-Verfahren: Proof of Work
- Blockchain-Anwendungen und -Typen
- Krypto-Währungen / Bitcoin

Quellen

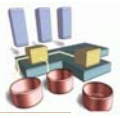
- Material basierend auf

- K.-U. Sattler.: VL Transaktionale Informationssysteme (TU Ilmenau)
- S. Maiyya et al.: Database and Distributed Computing Fundamentals of Blockchains, Tutorial @VLDB 2018, (https://tinyurl.com/w585fve)

Motivation

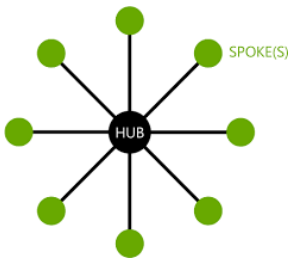
- Transaktion = technisches Konstrukt zur Sicherstellung von
  - Fehlersicherheit / Korrektheit bei Systemausfällen
  - Konsistenz / Korrektheit bei konkurrierenden Zugriffen

- Analogie Finanzwesen: Bau von Systemen
  - Repräsentation von Eigentümerschaft und -wechsel: Bankkonto, Finanzen, Grundstück, Reisebuchung, Warenlieferung …
  - Nachvollziehbarkeit aller Änderungen (Provenance)
  - Verhindern von doppelten Ausgaben („double spending“): das gleiche Produkt mehrfach verkaufen, das gleiche Geld zweimal ausgeben, …

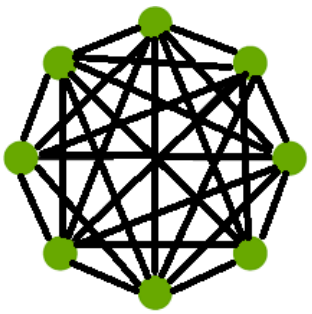


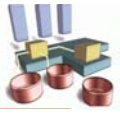
Mögliche Lösungen

- Lösung 1: zentralisiertes System (Bank) mit Kontrolle über Zustände (Eigentum) und Zustandswechsel
  - Nutzung eines zentralisierten DBS mit Transaktionskontrolle
  - auch bei Einsatz von parallelen/verteilten DBS bleibt zentrale Kontrolle bzw. begrenzte Knotenautonomie (Verteilungstransparenz)



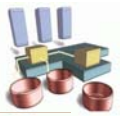
- Lösung 2: Blockchain-Systeme / verteilte Ledger-Systeme
  (*DLT: Distributed Ledger Technology*)
  - „Ledger“ (Buchführungssystem = Datenbank) zur Repräsentation des Zustandes
  - verteilte Kopien mit Append-Only Änderungen
  - globale Identität (Signatur) von Akteuren
  - Transaktion für Zustandsübergänge (Synchronisation gegen Double Spending)





Zielsetzungen

- mit Blockchain/DLT sollen mögliche Probleme zentraler Systeme lösen
- keine Abhängigkeit von „trusted third parties“
  - auch kein Vertrauen gegenüber anderen Teilnehmern erforderlich
- gleichberechtigter Zugriff auf Daten für alle Teilnehmer
- Daten können nicht manipuliert /gelöscht werden
- besserer Schutz gegenüber Angriffen
  - kein Single Point of Failure
- hohe Skalierbarkeit
- Nutzung in zahlreichen Anwendungen

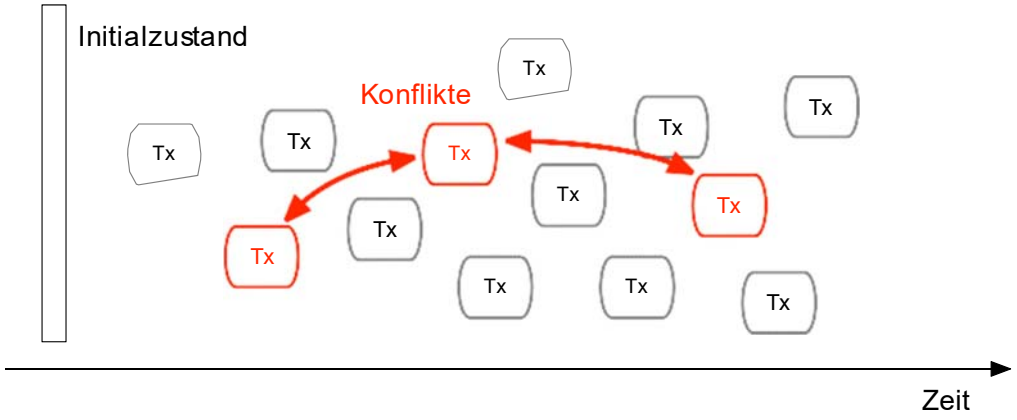


Blockchain: Begriff

- Blockchain = verteiltes System zur Verwaltung von Datensätzen mit dem Ziel, Konsens über den Zustand zu erzielen
- Eigenschaften
  - keine zentrale Instanz
  - Teilnehmer …
    - müssen andere Teilnehmer nicht kennen
    - müssen anderen Teilnehmern nicht vertrauen
    - können sich dennoch über einen Zustand einig sein
- Prinzip
  - Konsens über den initialen Zustand (z.B. leerer Zustand)
  - P2P-Netz aus Teilnehmern (Netzwerkknoten)
  - Transaktionen werden im Netzwerk angezeigt und weitergeleitet
  - Verhindern der Manipulation von Existenz oder Inhalt bereits ausgeführter Transaktionen

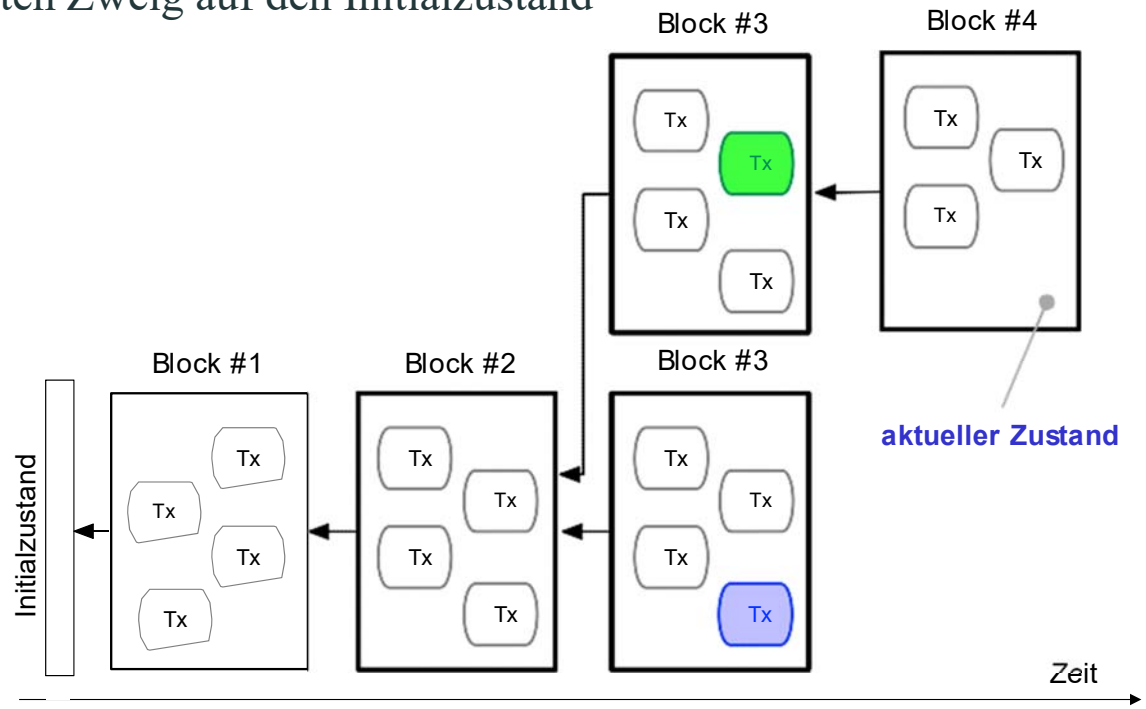
Probleme verteilter Transaktionen

- (kurzzeitig) unterschiedliche Zustände der Knoten (Konsistenzproblem)
- Reihenfolge der Transaktionen
- Versuche doppelter Ausgaben
- Konflikte zwischen Transaktionen bzw. Abhängigkeit von in Konflikt stehenden Transaktionen



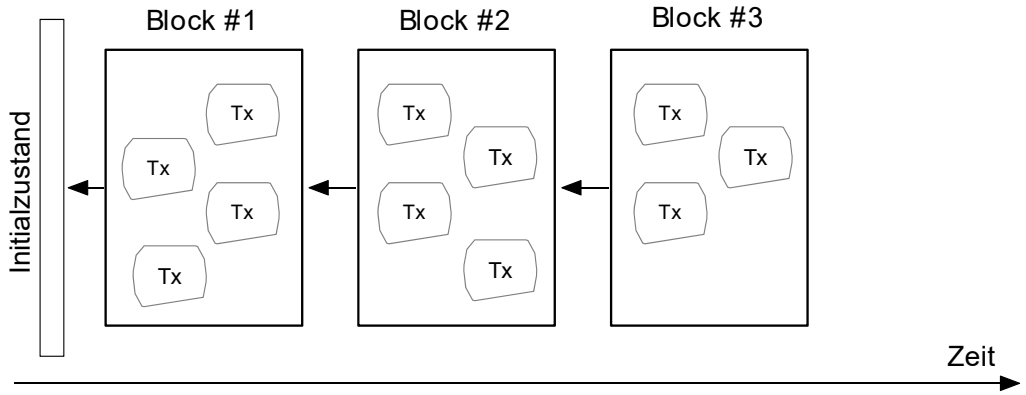
Blockchain-Elemente: Ledger

- Ledger = Blockkette (enthält Transaktions-Log)
  - jeder Knoten im Netzwerk verwaltet eigene Kopie des Ledgers
  - aktueller Zustand = Anwendung aller Transaktionen der Blöcke im längsten Zweig auf den Initialzustand



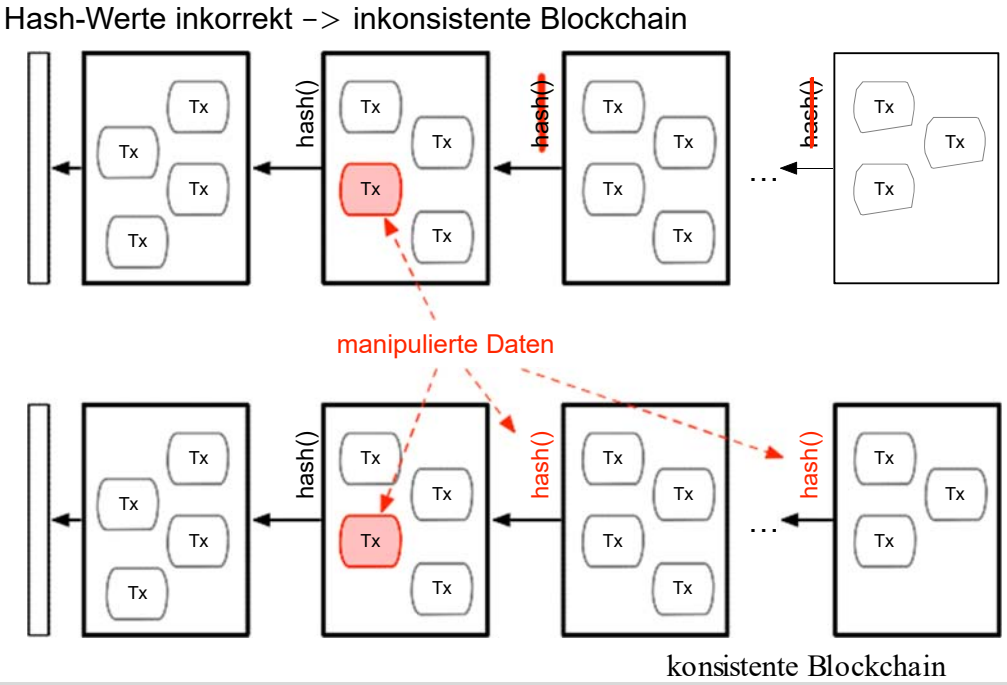
Blockchain-Elemente: Blöcke

- Lösung:
  - Zusammenfassung von Transaktionen zu Blöcken (Block-)
  - Blöcke werden verkettet (-chain), d.h. ein Block basiert auf seinen Vorgänger
    - Block enthält kryptographisch sicheren Hashwert seines Vorgängerblocks
- Blockinhalt: Transaktionsdaten, Zeitstempel, Hash des Vorgängerblocks (unveränderlich)
- jeder Teilnehmer kann jederzeit neuen Block erstellen



Blockchain: Manipulationssicherheit

- Verkettung der Blöcke durch Hash-Zeiger
  - Manipulation eines Blockinhaltes soll erkannt werden können
  - alleine nicht ausreichend: Ersetzen einer Teilkette durch eine manipulierte Teilkette muss extrem schwer gemacht werden

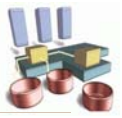


Blockchain: Ablauf

- Ausgangsbasis: vollständige Replikation des Ledgers

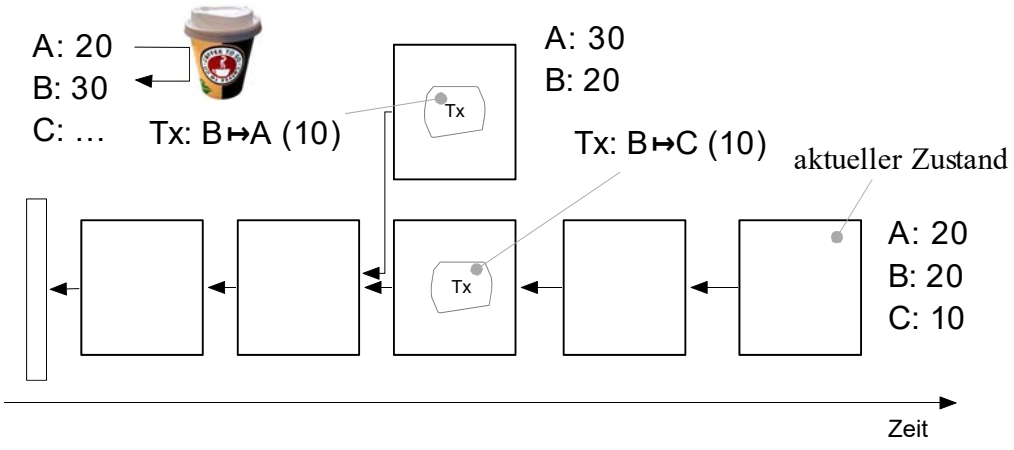
1. neue Transaktionen werden in einem Block gruppiert; Block hat feste Größe (z.B. 1 MB)
2. Netzwerkknoten validieren Transaktionen bzgl. Konflikten mit anderen Transaktionen des gleichen Blocks oder vorhergehender Blöcke
3. Knoten müssen sich darüber einig sein, welcher Block als nächster zur Blockchain hinzugefügt wird

-> Konsensverfahren



Double Spending in einer Blockchain

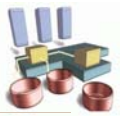
- Beispiel
  - B kauft bei A einen Becher Kaffee und zahlt dafür 10 Einheiten
  - B erzeugt weitere Transaktion, in der diese 10 Einheiten an C überwiesen werden sowie schnell hintereinander neue Blöcke
  - B behält Kaffee, A bekommt das Geld nicht



Konsens in der Blockchain

- Eignung bekannter Konsensverfahren wie Paxos …?
  - keine Behandlung byzantinischer Fehler (bösartiges Verhalten von Teilnehmern/Knoten, z.B. Austausch gefälschter Nachrichten)
  - Knoten müssen bekannt sein
  - auch in bekannten Verfahren für byzantinische Fehlertoleranz müssen Knoten bekannt sein

- anderer Ansatz notwendig ->Proof of Work (PoW)



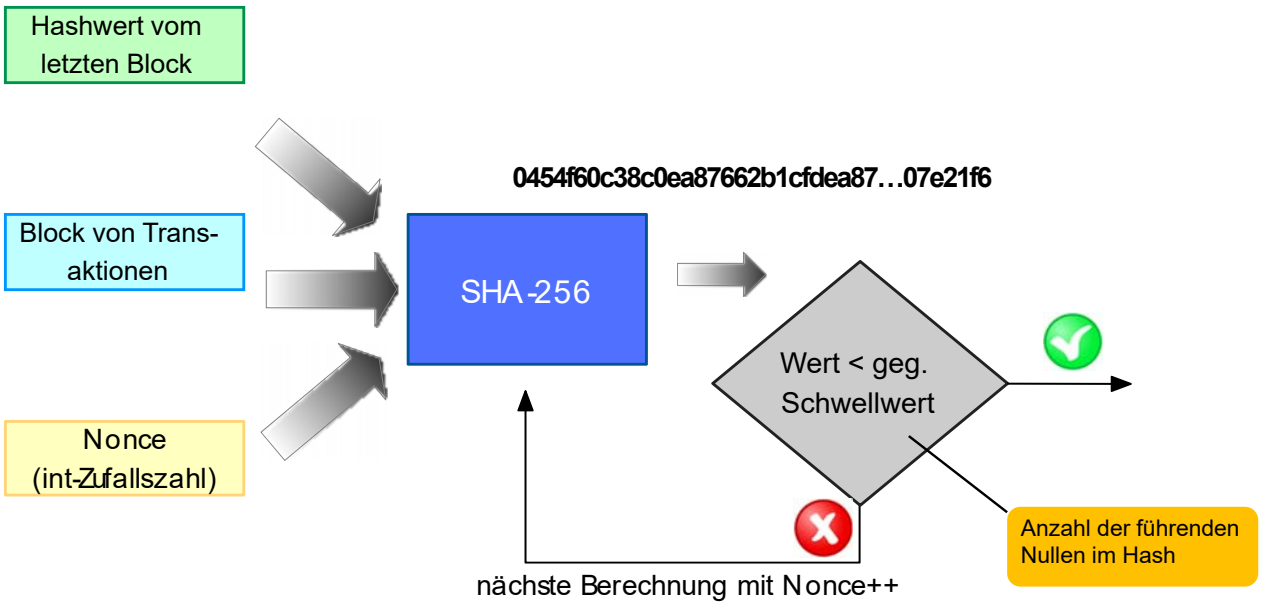
Proof of Work

- Erzeugung eines neuen Blocks soll aufwändig sein (in Zeit und Rechenkosten)

  -> Lösung einer Aufgabe (Proof of Work) = Mining
  - PoW muss validierbar sein
  - PoW ist Teil des neu erzeugten Blocks
  - nur Blöcke mit gültigem PoW sind gültige Blöcke
  - Beispiel Bitcoin: *Hashcash*

- PoW Anforderungen
  - aufwändige Berechnung (nur mit Brute Force) aber einfache und schnelle Validierung
  - muss abhängig vom zu erzeugenden Block sein (Vermeidung von Vorabberechnungsangriffen)
  - variabler Schwierigkeitsgrad
  - Anreizsystem: Mining selbst sollte lohnenswert sein: Belohnungstransaktion (Reward Transaction)
  - Teil des neuen Blocks (*Coinbase* in Bitcoin)

PoW: Hashcash

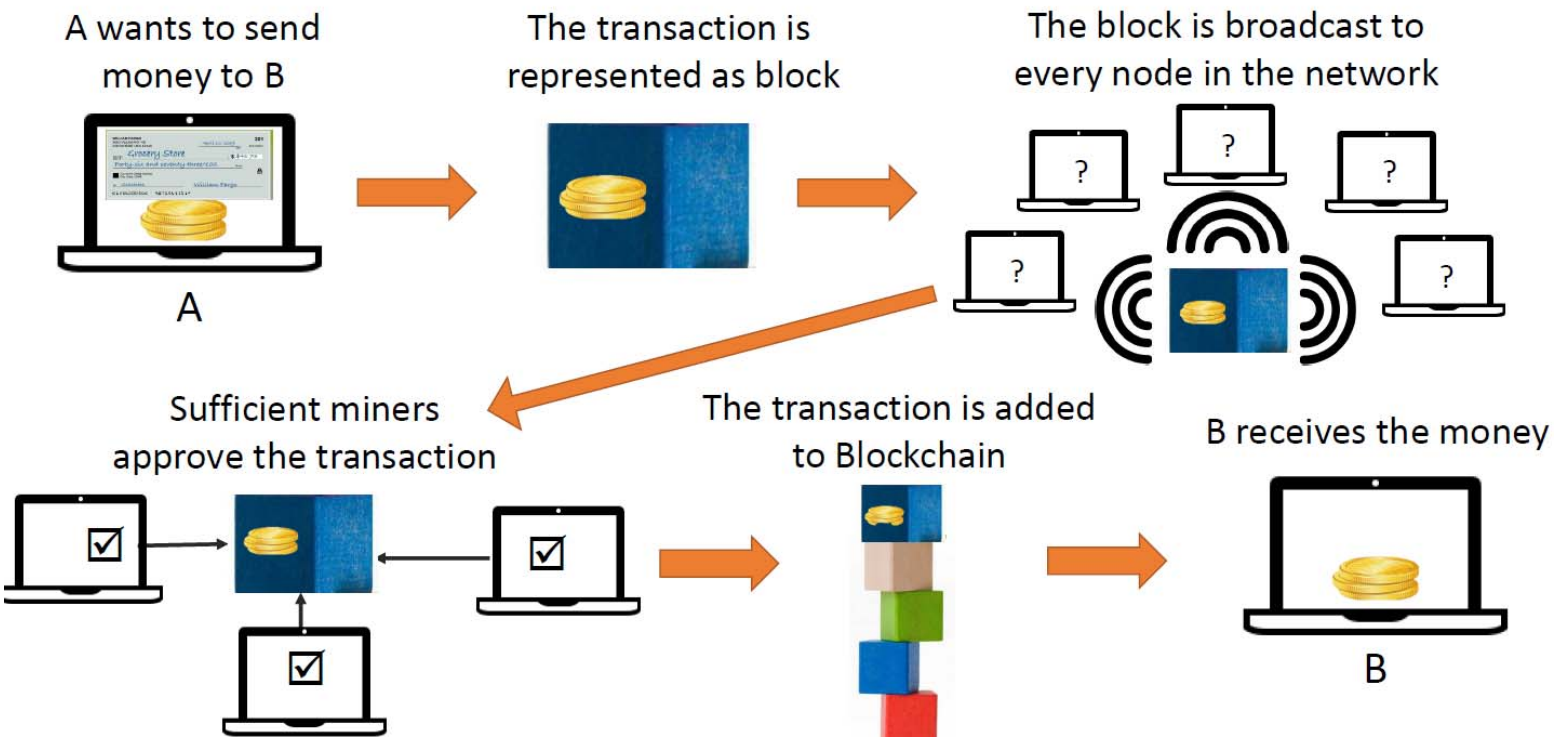


PoW: Ablauf

- wenn Knoten Aufgabe gelöst hat (Mining abgeschlossen):
  - füge Block von Transaktionen der Blockchain hinzu
  - Multicast (Flooding) der Lösung an andere Netzwerkknoten
  - Netzwerkknoten validieren und akzeptieren Lösung

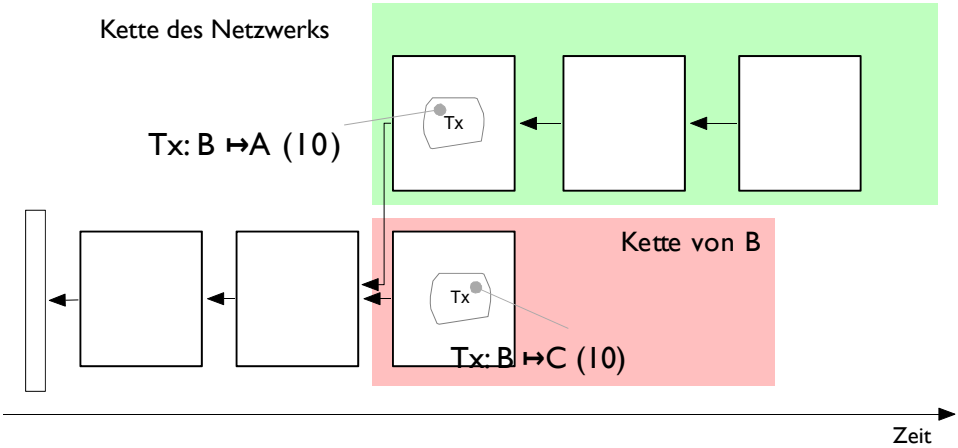
- in welchem Zweig der Blockchain sollte ein Miner arbeiten?
  - für Belohnung: Zweig muss Teil des aktuellen Zustands sein (=längster Zweig)
  - keine Koordination notwendig!

Gesamtablauf



Double Spending bei PoW

- Beispiel:
  - B erzeugt wieder 2 Transaktionen in verschiedenen Blöcken, nur die erfolgreich validierte wird veröffentlicht
  - Zuordnung weiterer Transaktionen zum Originalblock, nur B arbeitet auf seinem Block
  - Netzwerk hat mehr Rechenleistung und kann schneller neue Blöcke erzeugen -> Kette von B bleibt kürzer



Netzwerkangriffe: 51%-Angriff

- Mehrheitsangriff: Angreifer kontrolliert über die Hälfte der Rechenleistung des Netzwerks
  - ermöglicht Double Spends, Rückgängigmachen von Transaktionen
  - Prinzip: eigene Blöcke schneller anlegen als der Rest des Netzes und nachträglich gültig machen
- nach Wikipedia:
  - 2014: Mining Pool GHash überschreitet kurzzeitig 50%-Marke
  - Attacken auf Bitcoin Gold (2018) und Ethereum Classic (2019)
- Gegenmaßnahmen
  - 51%-Angriff ist ein auffälliges Ereignis -> z.B. Hard Forks in Bitcoin
    - ersten Block einer verdächtigen Kette ungültig erklären
  - eingebaute Anreizmechanismen:
    - hohe Miningkosten im Falle einer Abwehr verloren
    - Belohnung pro Block kann nicht erhöht werden
    - Senden anderer Transaktionen kann nicht verhindert werden

Blockchain: Anwendungen

- Krypto-Währungen (Bitcoin, Etherum etc.)
- Auditing: Aufzeichnung sicherheitskritischer Operationen
  - Zugriff auf bzw. Veränderung von Ressourcen (z.B. Daten, Dokumente)
  - Zugriff auf Gesundheitsakten, …
- Smart Contracts: Protokoll zur Repräsentation, Verifikation und (teilweisen) Ausführung von Verträgen, z.B. Ausführung von Prozeduren bei bestimmten Transaktionen
- dezentrale Energieversorgung und –Abrechnung
- Lieferketten: Dokumentation der Teilschritte

  …

- generell: Verzicht auf zentrale Instanz

Blockchain/DLT-Typen

- öffentliche vs. private Blockchains
- öffentliche Blockchains (z.B. für Kryptowährungen)
  - lassen beliebige Teilnehmer zu
  - maximale Transparenz
  - keine Vertraulichlichkeit (privacy) und sehr hoher Ressourcenbedarf
- private Blockchains (z.B. für Unternehmensanwendungen)
  - durch Eigentümer oder Konsortium kontrollierter Teilnehmerkreis (*permissioned* blockhains)
  - effizientere und kostengünstigere Realisierung von Transaktionen
  - Beispielrealisierung: Hyperledger (www.hyperledger.org)

Anwendung: Krypto-Währungen

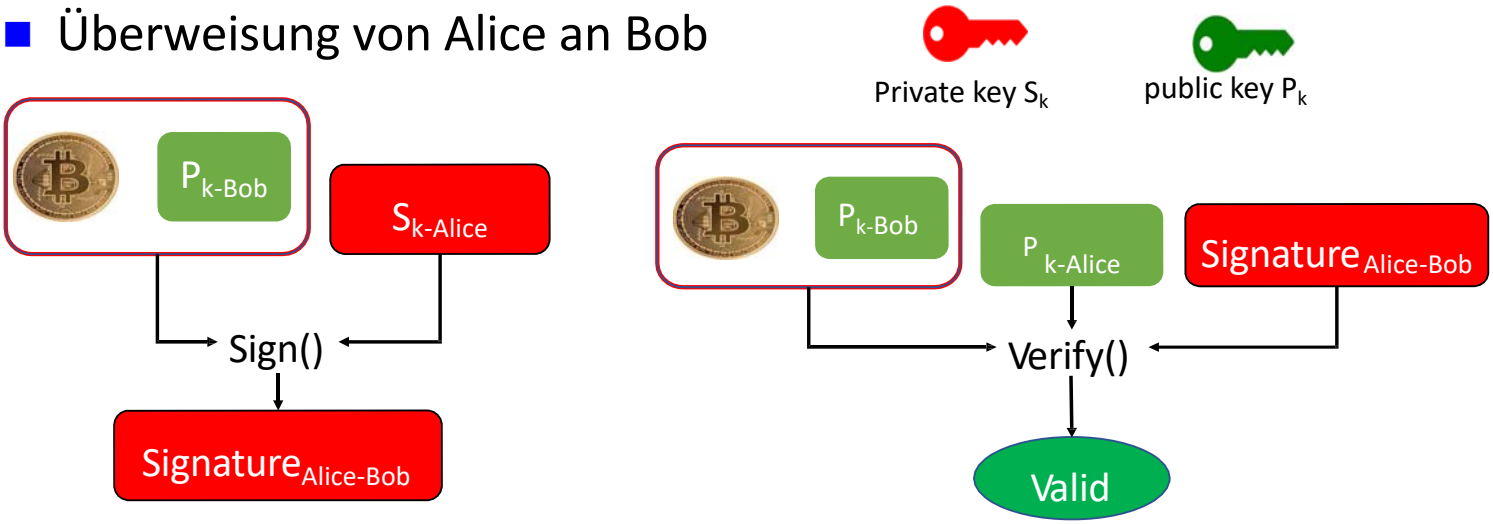
- digitale Zahlungsmittel
- Historie: Bitcoin
  - 2008: Artikel von „Satoshi Nakamoto" *Bitcoin – A Peer-to-Peer Electronic Cash System* https://bitcoin.org/bitcoin.pdf
  - 2009 Open-Source-Software, eigentlicher Start
  - starke Kursschwankungen
    - all-time high (Dec. 2017): 19,78 TE, Jan. 2020: 8 TE pro bitcoin
- aktuell weit über 1000 Währungen
  - 80% aller Initial Coin Offerings mit betrügerischem Hintergrund (Wikipedia)
  - weniger als die Hälfte überlebt ersten vier Monate
- Fiatgeld:
  - Zahlungsmittel ohne inneren Wert (im Gegensatz zu Warengeld wie Gold)
  - benötigt Vertrauen der Beteiligten, z.B. Zentralbank, Staat, sich gegenseitig kontrollierende Teilnehmer

Krypto-Währungen: Bausteine

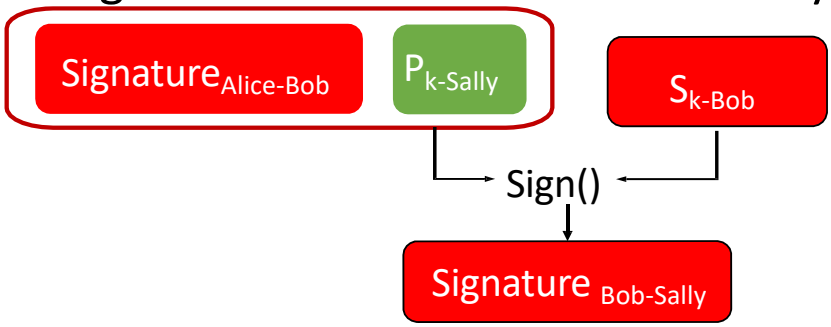
- Vernetzung der Teilnehmer: P2P-Netz statt zentrale Instanz
  - mehr als 10.000 Bitcoin-Knoten nach https://bitnodes.earn.com
- kryptographische Signaturen: Public-Key-Kryptosystem
  - öffentlicher Schlüssel = Kontonummer
  - privater Schlüssel = Verfügungsgewalt über Konto
  - Überweisung: Betrag + öffentlicher Schlüssel des Empfängerkontos, signiert mit privatem Schlüssel des Senders (=Transaktion)
  - Überweisung wird im Netz verteilt und kann von allen überprüft werden
- Buchführung: Transaktionen werden im Ledger voll repliziert auf allen Knoten verwaltet
- Bitcoin-Transaktionen
  - keine expliziten Konten: Guthaben = eingegangene Gutschriften, die noch nicht weiter überwiesen wurden

Digitale Signaturen und Bitcoin

- Überweisung von Alice an Bob



- Weitergabe der Bitcoins von Bob an Sally



Hashing H(x)

- Kombination von Signaturen und Public Keys über Hashing
  - Eingabe: String beliebiger Länge
  - Ausgabe fester Länge (z.B. 256 Bits)
  - effizient berechenbar
- Bitcoin nutzt SHA-256 (**S**ecure **H**ash **A**lgorithm)

SHA256 ($Signature_{Alice\text{-}Bob}$ || $P_{k\text{-}Sally}$) =

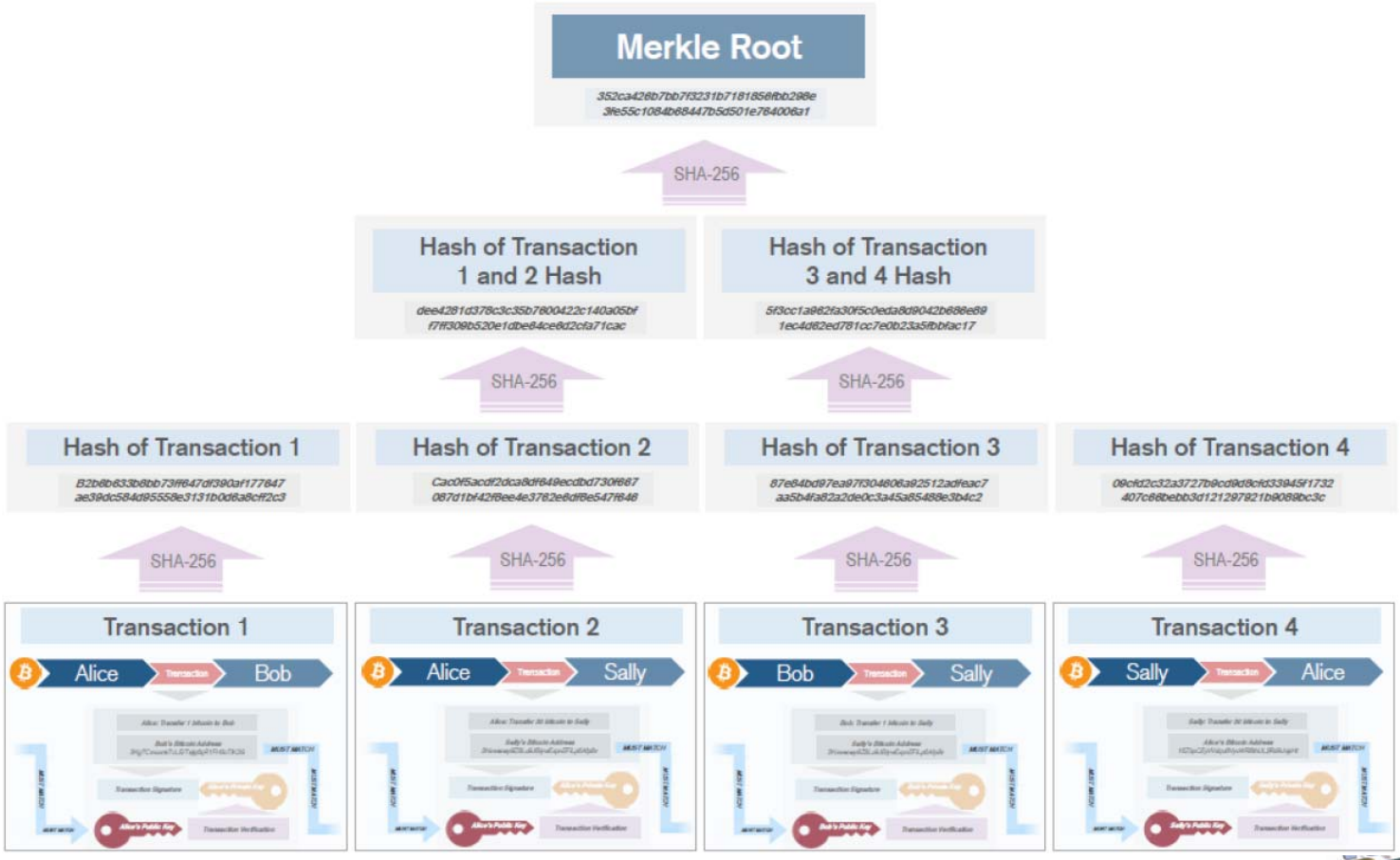
256-bit (32-byte) eindeutiger String

- Eigenschaften:
  - *kollisionsfrei*: keine zwei x, y so dass H(x) = H(y)
  - *sicher:* unmöglich x aus H(x) abzuleiten (one-way hash function)

Bitcoin: Transaktionen und Blöcke

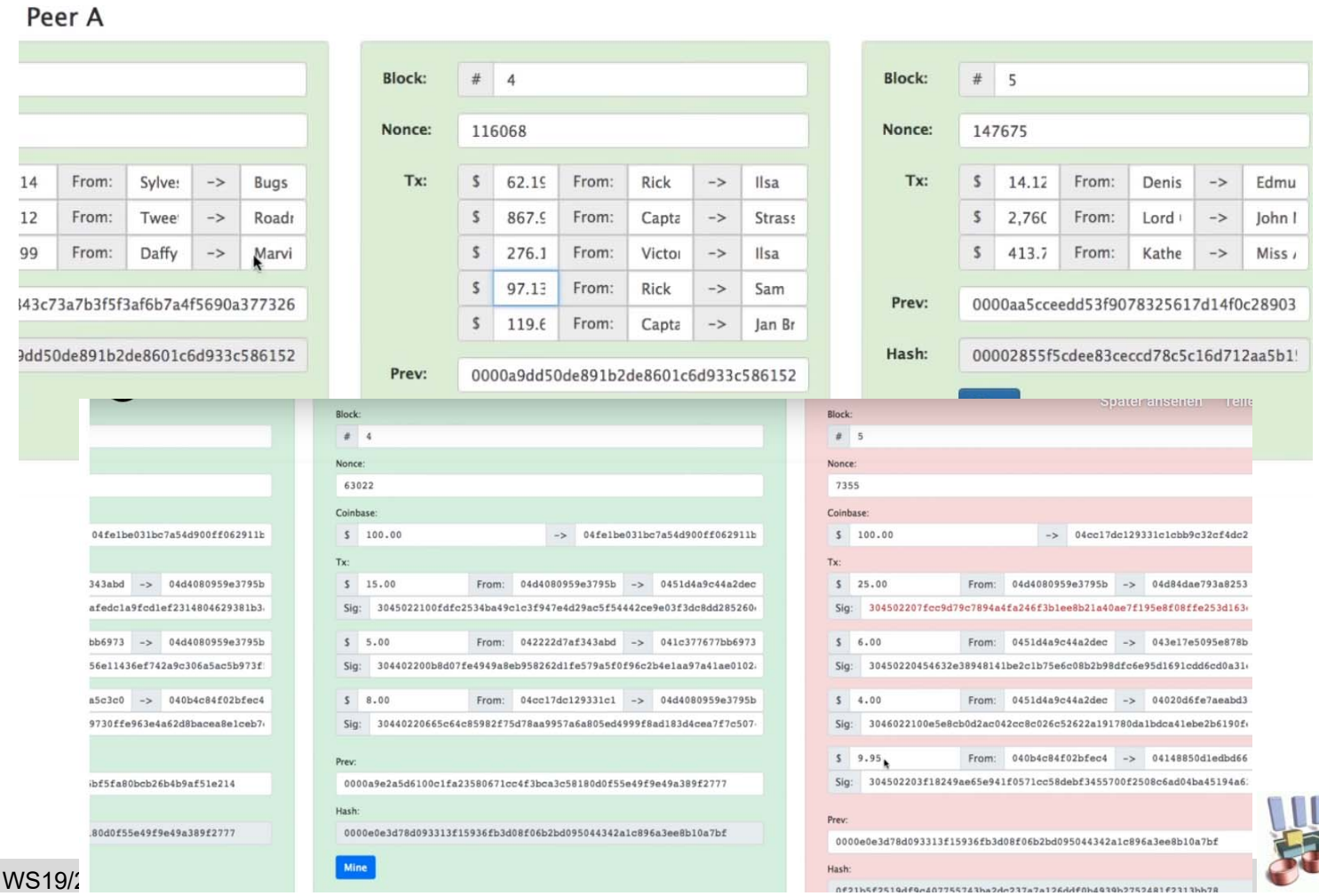
- Transaktionen
  - Inhalt: Senderadresse, Empfängeradresse, Betrag, Signatur
  - selbstgewählte Transaktionsgebühren
  - mit privatem Schlüssel des Senders signiert
  - im Netzwerk validiert und verbreitet
- Blöcke
  - 1. Block = Genesis-Block
  - neue Blöcke werden durch Mining erzeugt (Übernahme von noch unbestätigten Transaktionen aus Mempool)
  - erste Transaktion eines Blocks enthält Überweisung neu erzeugter Bitcoins für Mining + Transaktionsgebühren (Reward)
  - Hashwert = paarweises Hashing der Transaktionen in *Merkle-Baum*, Hashwert des Wurzelknotens (Root-Hash) als Prüfsumme des Blocks
- Mining: PoW wie beschrieben (Nonce-Variation, Schwellwert)
  - per Flooding an alle Bitcoin-Knoten verteilt
  - jeder Knoten muss selbst bestimmen, welcher Block/welche Kette gültig ist (Konsens)
  - Miner erhält geschürfte Bitcoins + Gebühren der enthaltenen Transaktionen

Merkle-Baum (Bsp.)



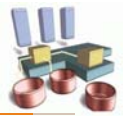
https://www.p-ic.at/uploads/simplex/images/BMVIT Dossier Blockchain 2017 FINAL.pdf

Demo https://andersbrownworth.com/blockchain/



Bitcoin: Fakten (Wikipedia)

- Stand 2018: Blockchain 190 GB
  - aktuell: 12,5 neu erzeugte Bitcoins pro Block; halbiert sich alle 4 Jahre
  - Transaktionskosten: 1.000 Satoshi (= 10 µBTC)
- max. 7 Transaktionen pro Sekunde (schlechte Skalierbarkeit)
- extremer Ressourcen/Energie-Bedarf
  - pro Bitcoin: 8 bis 13 Tonnen CO2-Ausstoß; etwa 42.000 kWh Strom
  - pro Transaktion: 300 kWh; Kreditkartentransaktion: 1-2 Wh
  - Stand Januar 2019: 47 Mrd. kWh/Jahr gesamt
  - ca. 75% aller Bitcoins in China geschürft (2017) – durch Kohlestrom aus der Inneren Mongolei



Zusammenfassung

- Blockchains/Distributed Ledgers: neues Paradigma für verteilte Daten- und Transaktionsverwaltung
- Popularität durch Kryptowährungen wie Bitcoin
- wesentliche Vorteile
  - gleichberechtigte Datennutzung, keine Abhängigkeit von zentralen Institutionen, keine Veränderung bereits erfolgter Transaktionen …
- technische Realisierung
  - Blockbildung und Verkettung durch Hashing verschlüsselter und signierter Transaktionen
  - vollständige Replikation der Blockchain
  - Validierung durch Mining und Konsensbildung
- zunehmende Nutzung über Kryptowährungen hinaus
- private Blockchains für Unternehmensanwendungen mit besserer Leistungsfähigkeit und geringerem Ressourcenbedarf