Merkle-Bäume: Was sie sind und wo sie benutzt werden

Thomas Braun

byte physics e. K.

12. April 2023

Intro

- Benannt nach Ralph Merkle, einem der Pioniere asymmetrischer Kryptographie[1]
- Auch bekannt als Hash-Bäume

[1] https://de.wikipedia.org/wiki/Ralph_Merkle

Das zu lösende Problem (1988)[2]

- Erzeugen digitaler Signaturen ohne (damals) aufwendige mathematische Operationen wie Modulo-Arithmetik
- Keine Begrenzung der Anzahl der Nachrichten
- Keine Explosion der Anzahl der Signaturen (es werden nur $\log_2(n)$ Signaturen für n Nachrichten benötigt)

---

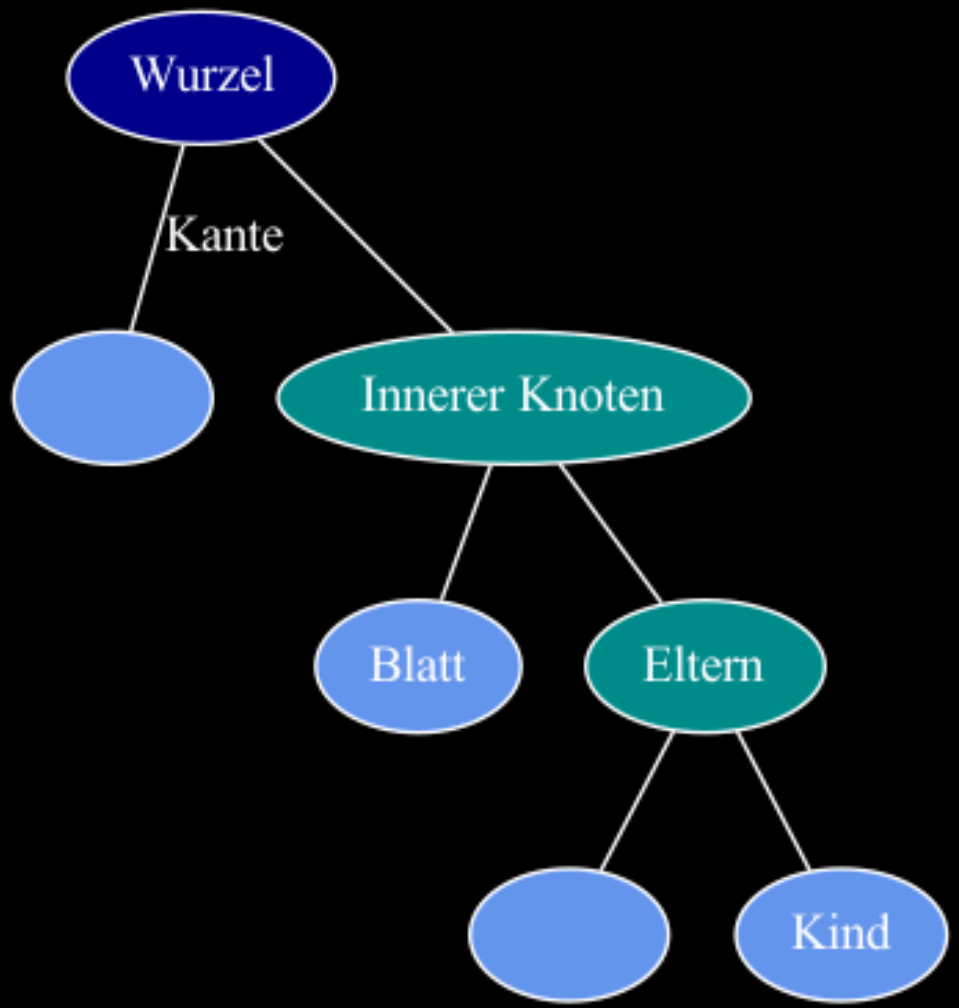
[2]Merkle, R.C. (1988). A Digital Signature Based on a Conventional Encryption Function. In: Pomerance, C. (Hrsg.) Advances in Cryptology — CRYPTO '87. Lecture Notes in Computer Science, vol 293. Springer. https://doi.org/10.1007/3-540-48184-2_32

Einschub: Bäume

- Abstrakter Datentyp in der Informatik
- Definition[3]: Ein Baum besteht aus einer Menge von Knoten und Kanten die besondere Eigenschaften aufweisen:
  - Jeder nicht leere Baum besitzt einen ausgezeichneten Knoten, die Wurzel
  - Jeder Knoten, außer der Wurzel, ist durch genau eine Kante mit seinem Elternknoten verbunden. Er wird Kindknoten genannt.
  - Ein Knoten ohne Kinder heißt Blatt.

[3] https://services.informatik.hs-mannheim.de/~schramm/ads/files/Kapitel10_01.pdf

Beispiel eines Baums

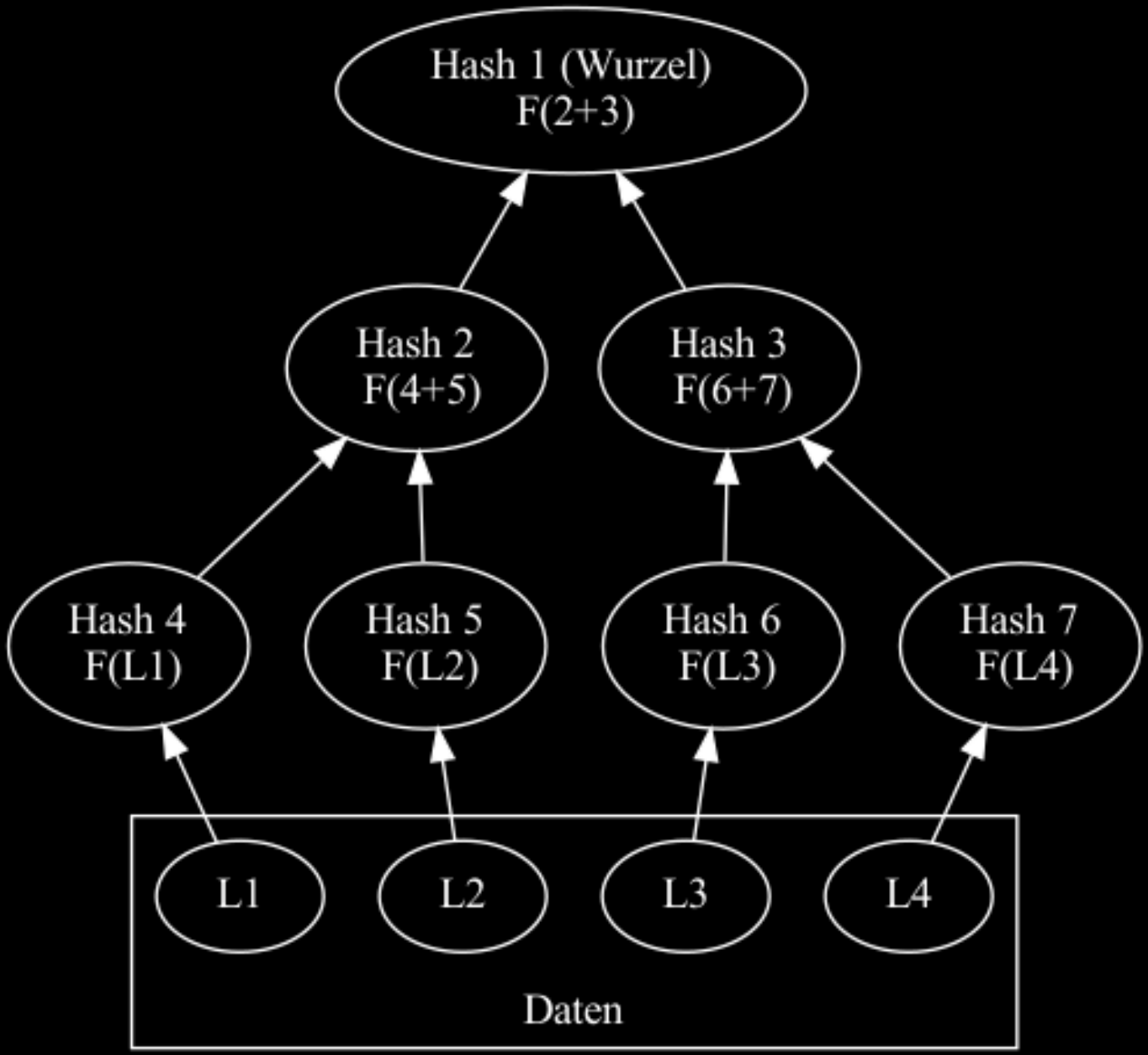


Einschub: Einwegfunktionen

- Mathematische Funktion $F(x) = y$ die leicht zu berechnen ist
- Deren Umkehrung $F^{-1}(y) = x$ aber sehr aufwendig zu berechnen ist
- Beispiele für kryptografische Einwegfunktionen: SHA-2/SHA-3/Whirpool[4]

[4] Diese Einwegfunktionen werden, Stand heute, als quantensicher betrachtet.

Binärer Hash-Baum



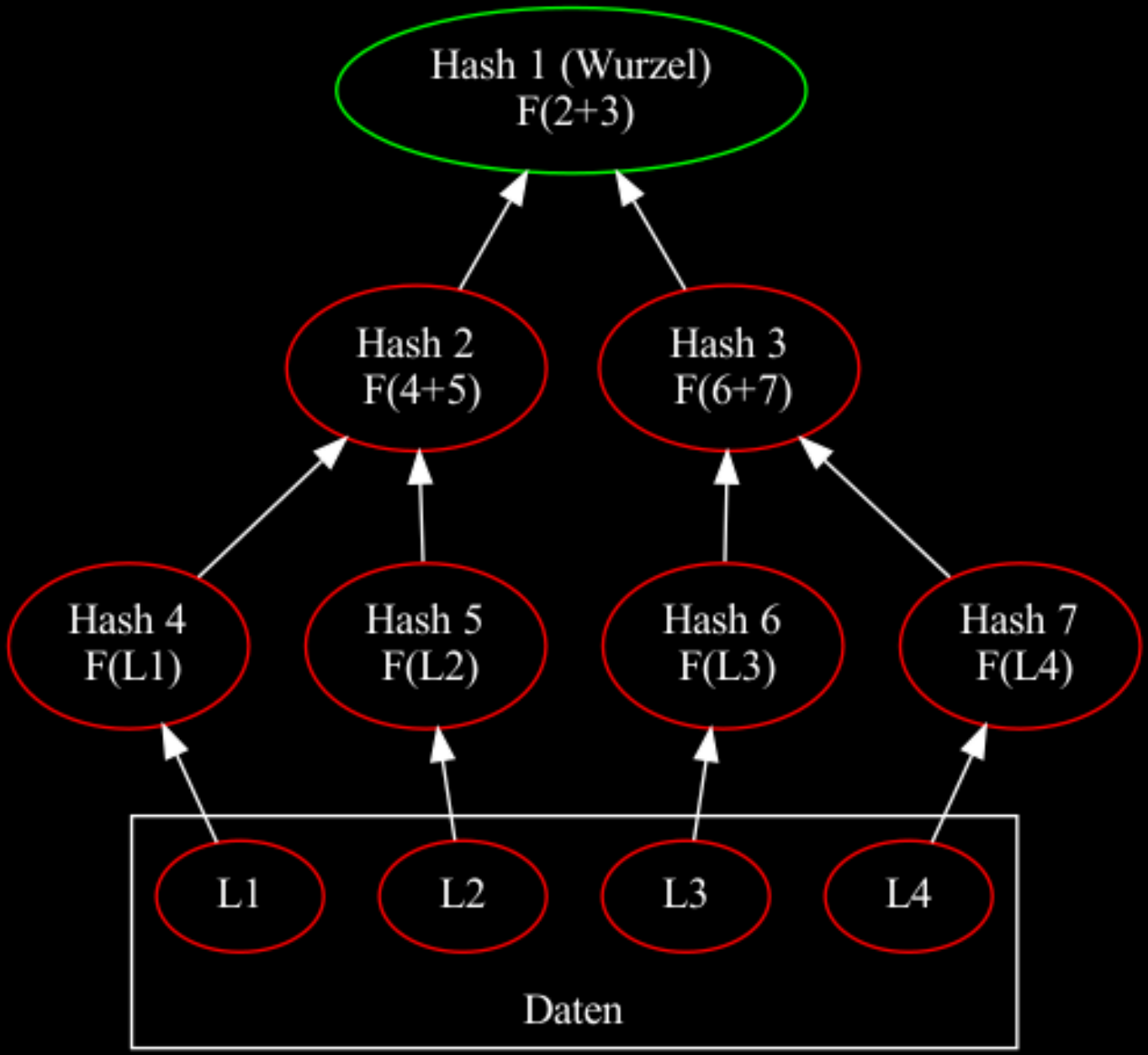
Beispiel: Tree Hash EXchange format (THEX)[5]

- Zweck: Datenintegrität sicherstellen beim Übertragen großer Datenmengen
- Nur der Wurzel-Hash muss über einen vertrauenswürdigen Kanal transportiert werden
- Sowohl Daten (L1, L2, L3, ...) als auch Innere-Hashes (H2, H6, ...) können von nicht vertrauenswürdigen Quellen stammen
- Paralleles Herunterladen und Überprüfen von Dateiabschnitten möglich

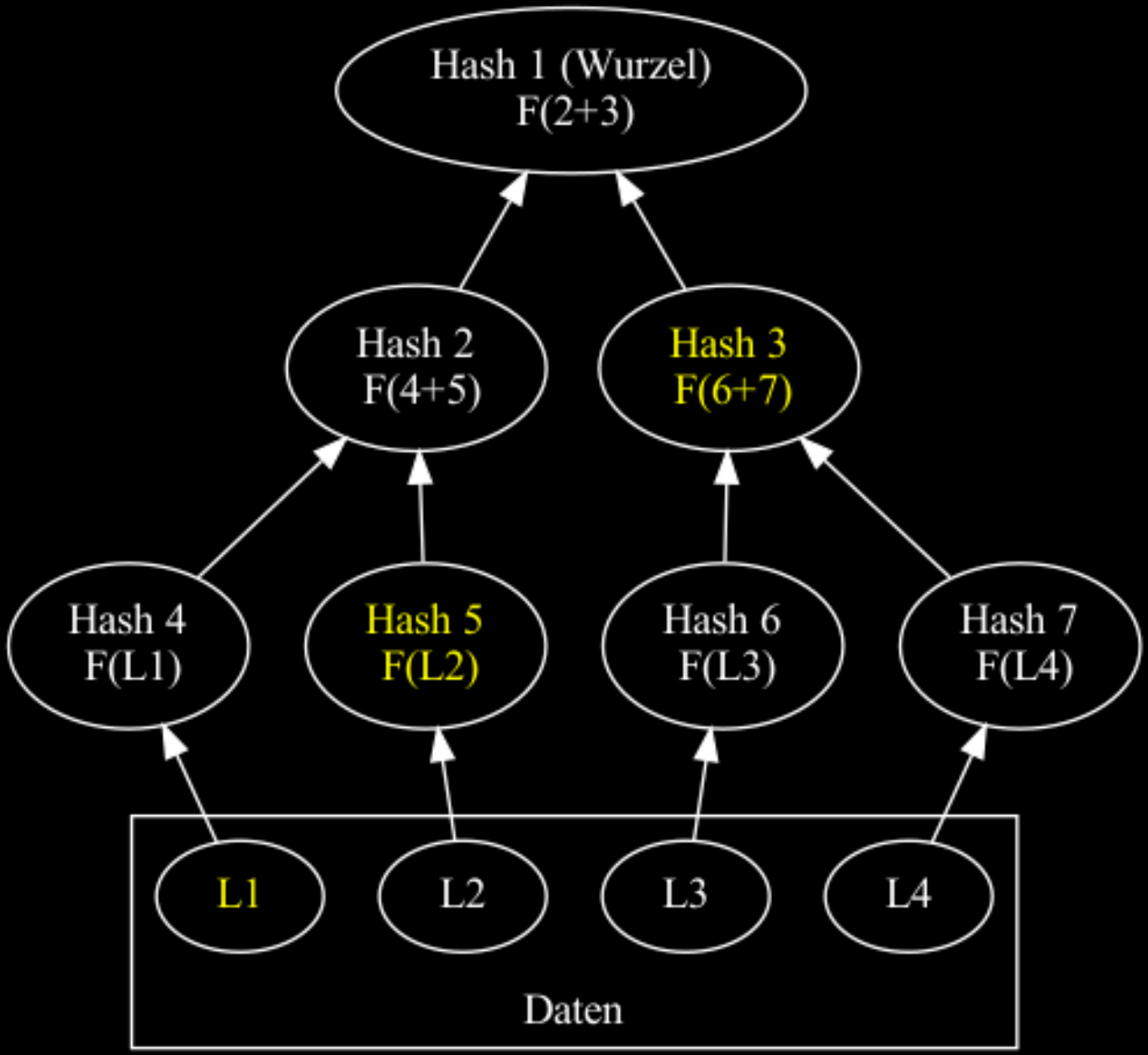
---

[5] `https://web.archive.org/web/20080316033726/http://www.open-content.net/specs/draft-jchapweske-thex-02.html`

Vetrauensanforderungen



Überprüfung von L1



Beispiel: Das Git Objektmodell[8]

- Alle Objekte in git (blob, tree, commit, tag) werden durch SHA-1/SHA-256[6]-Hashes identifiziert.
- Durch Verweise
  - tree: (tree, blob)
  - commit: (tree, commit)
  - tag: (commit)

  ensteht ein Baum.[7]

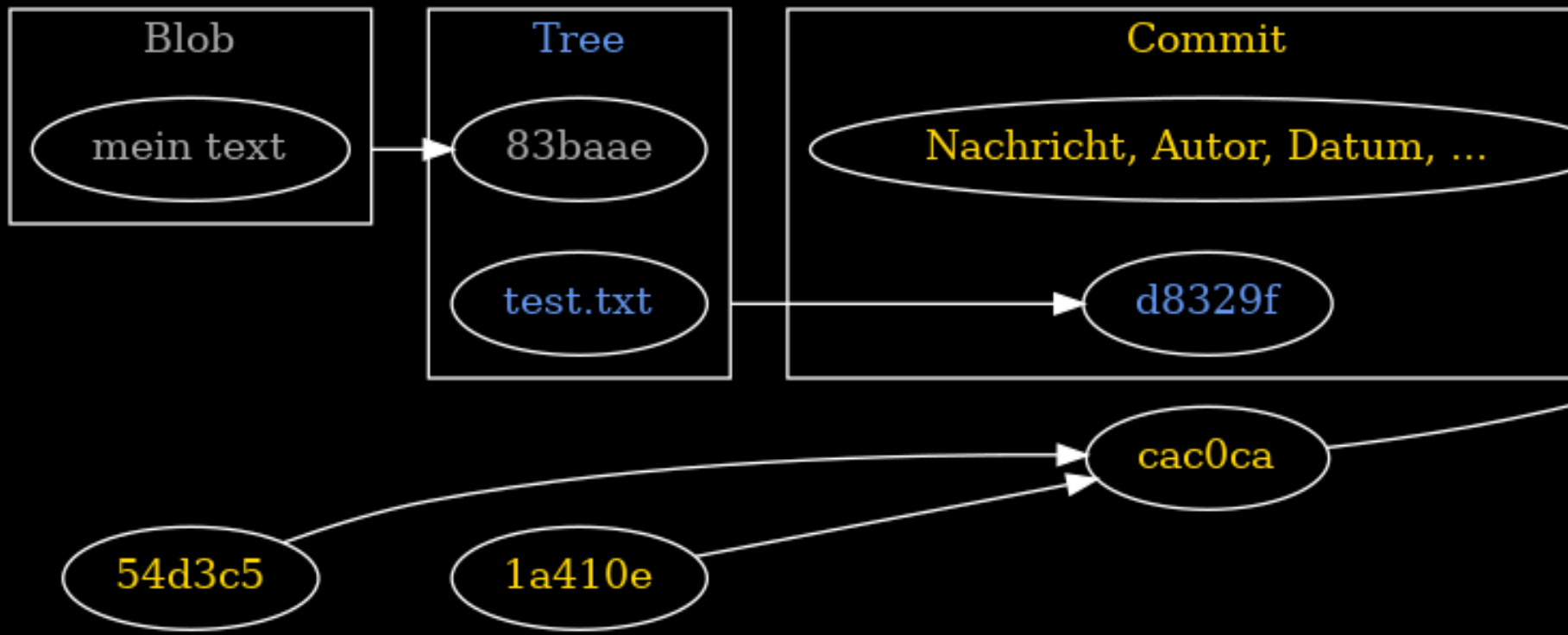
---

[6] `https://lwn.net/Articles/898522`

[7] Das sind nur die üblicherweise betrachteten Verweise, auch andere Kombinationen sind möglich.

[8] `https://git-scm.com/book/en/v2/Git-Internals-Git-Objects`

Git-Repository als Hash-Baum



Bitcoin

- Digitales Geld, Paper[9] und Quellcode öffentlich seit 2009[10]
- Erfordert keine zentrale und vertrauenswürdige Instanz
- Alle Transaktionen sind öffentlich
- P2P-Infrastruktur für Verwalten der Transaktionen
- Einzige Anforderung: Die Mehrheit der Teilnehmer, in Einheiten von Rechenpower, muss vertrauenswürdig sein

---

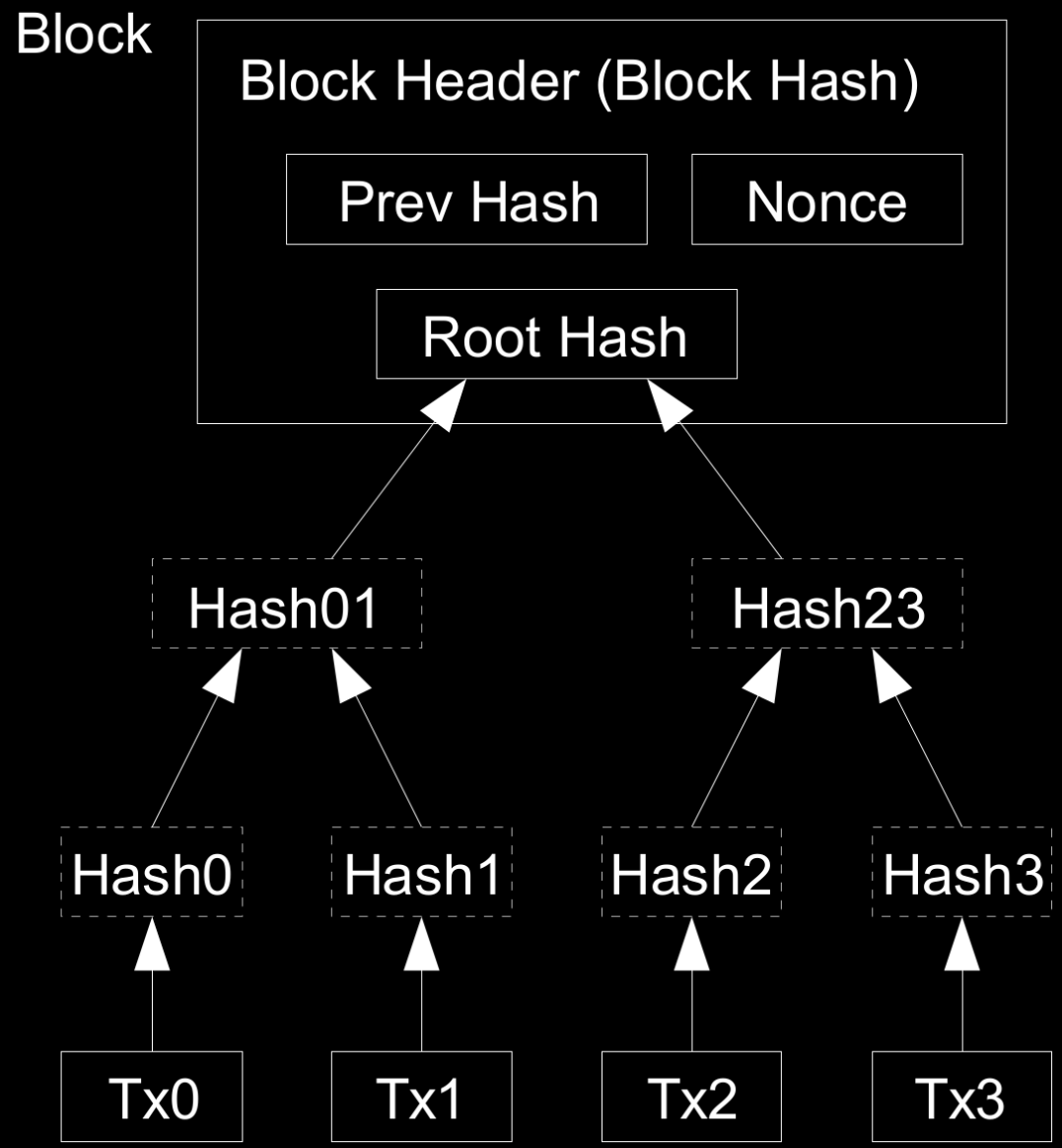
[9] https://bitcoin.org/bitcoin.pdf

[10] https://github.com/bitcoin/bitcoin/commit/e071a3f6

Organisation der Transaktionen

- Eine Menge X an Transaktionen werden gesammelt und von einem Teilnehmer des Netzwerks mittels dem Lösen eines Rätsels (Proof-of-Work) beglaubigt.
- Als Belohnung für diese Arbeit darf jeder Teilnehmer in der ersten Transaktionen jedes neuen Blocks einen Bitcoin erzeugen und sich selbst überweisen.
- Zur effizienten Speicherung der Transaktionen wird ein Merkle-Baum verwendet

Merkle-Baum der Transaktionen



Fragen ?