What Is A Blockchain

- A **decentralized** **computation and information sharing platform** that enables **multiple authoritative domains**, who **do not trust** each other, to **cooperate, coordinate** and **collaborate** in a **rational decision making process**

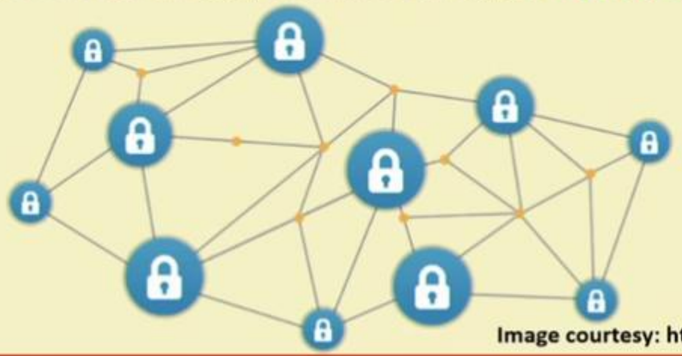
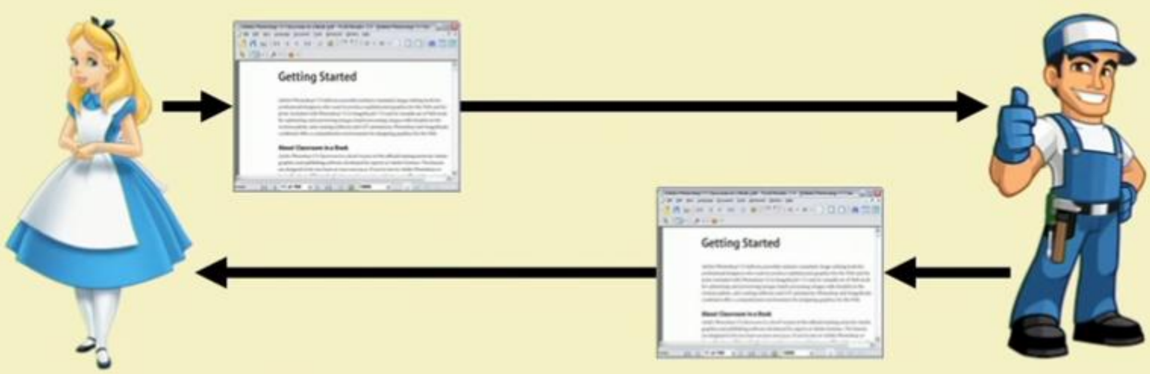


Image courtesy: https://blog.exchangeunion.com

Microsoft Word to Google Doc – Sharing Information

- Traditional way of sharing documents



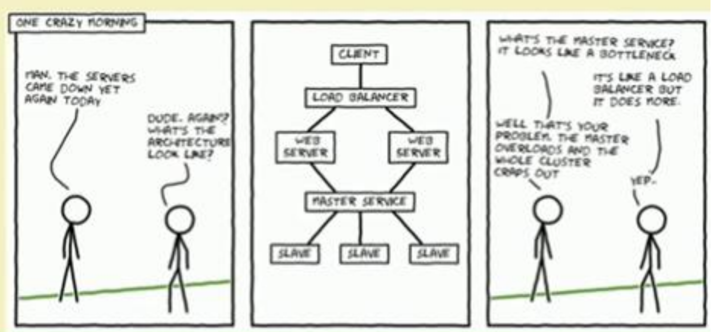
Microsoft Word to Google Doc – Sharing Information

- Shared Google doc – both the users can edit simultaneously

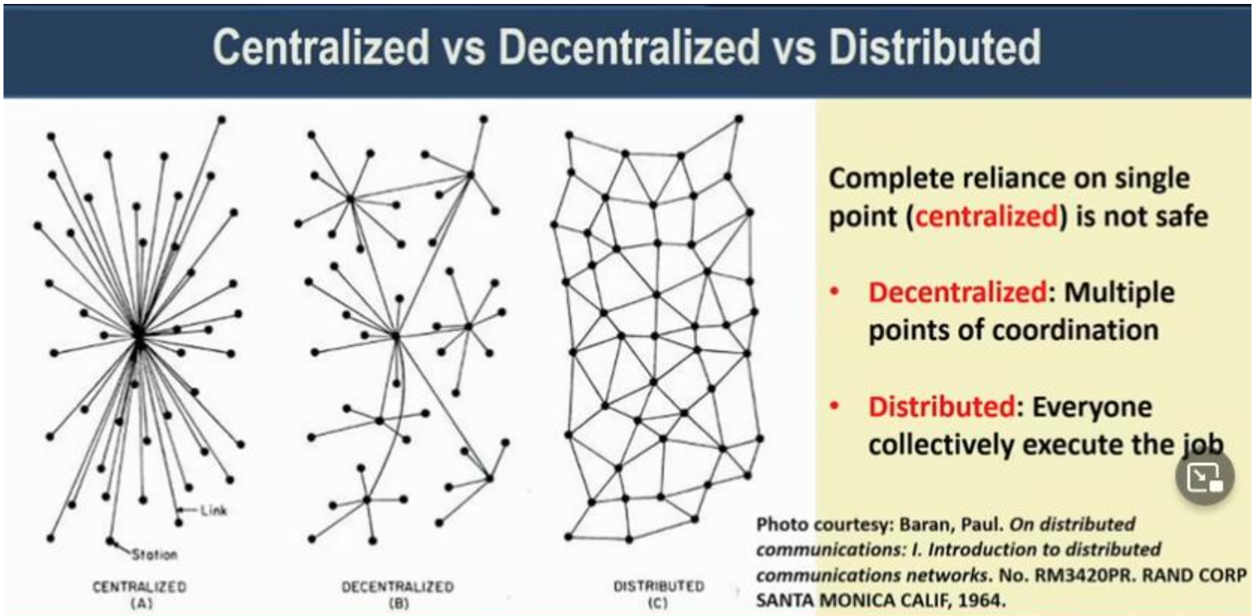
The environment is still centralized.
Does centralized system harm?

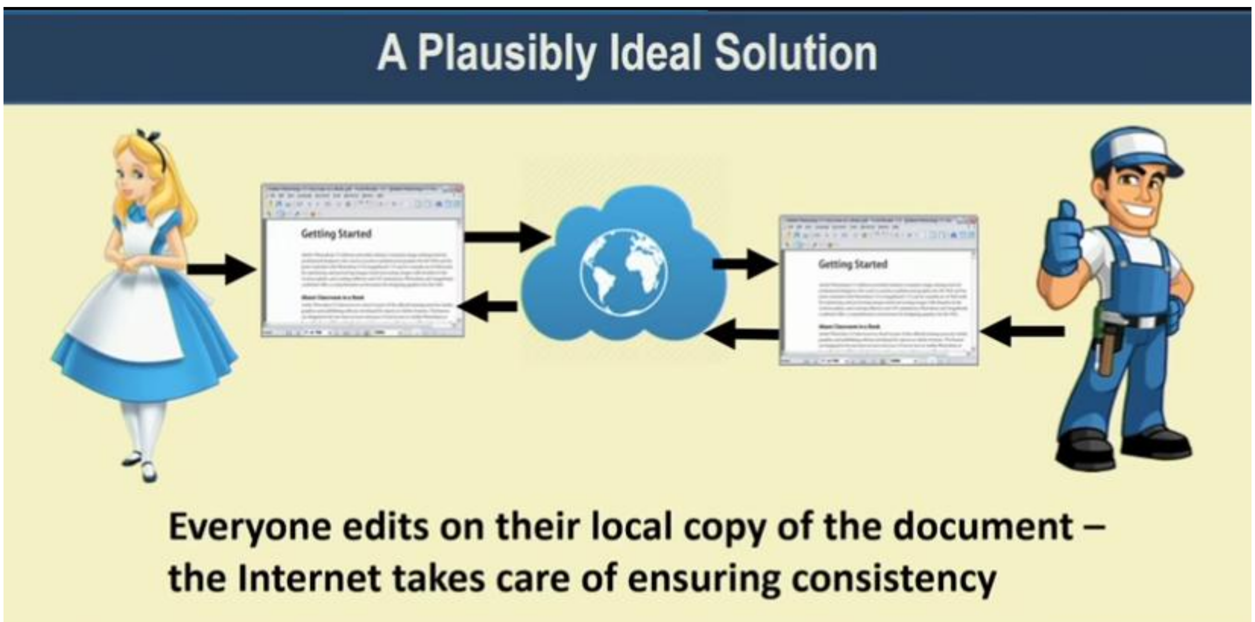
Problems with a Centralized System

- **A single point of failure**
  - If you do not have sufficient bandwidth to load Google doc, you'll not be able to edit
  - What if the server crashes?

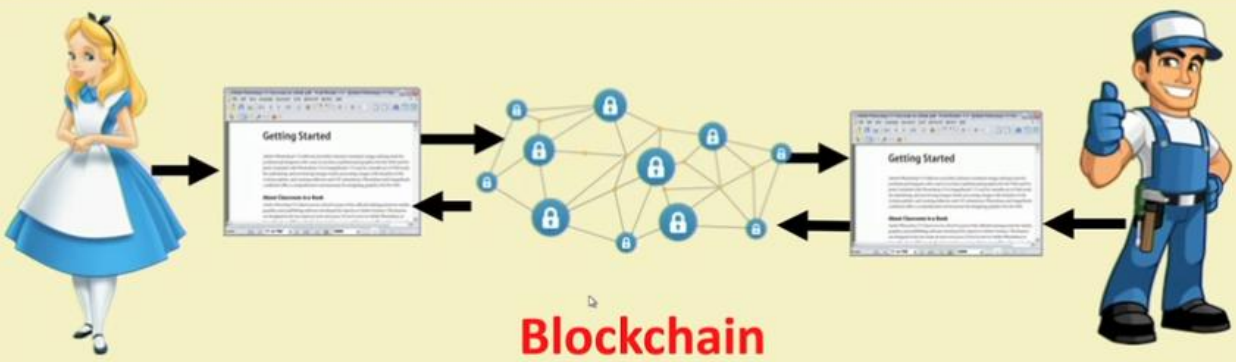


mage courtesy: http://timkellogg.me/


Centralized vs Decentralized vs Distributed
Complete reliance on single point (centralized) is not safe
• Decentralized: Multiple points of coordination
• Distributed: Everyone collectively execute the job
Link
Station
CENTRALIZED (A)
DECENTRALIZED (B)
DISTRIBUTED (C)
Photo courtesy: Baran, Paul. On distributed communications: I. Introduction to distributed communications networks. No. RM3420PR. RAND CORP SANTA MONICA CALIF, 1964.


A Plausibly Ideal Solution
Getting Started
Getting Started
Everyone edits on their local copy of the document – the Internet takes care of ensuring consistency

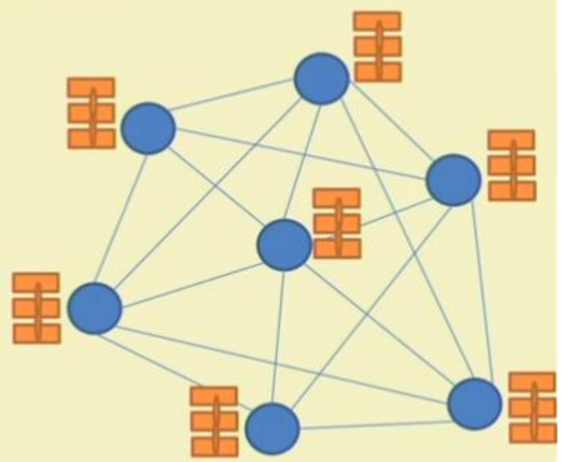
Blockchain – The Internet Database to Support Decentralization



A decentralized database with strong consistency support

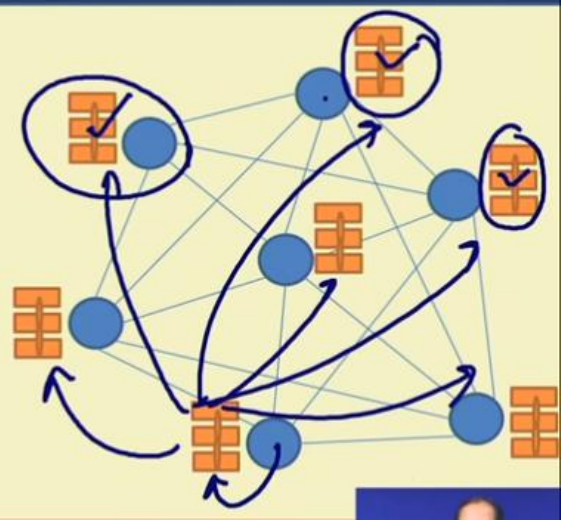
A Very Simplified Look of the Blockchain

- Every node maintains **a local copy** of **the global data-sheet**
- The system ensures consistency among the local copies
  - *The local copies at every node is identical*
  - *The local copies are always updated based on the global information*

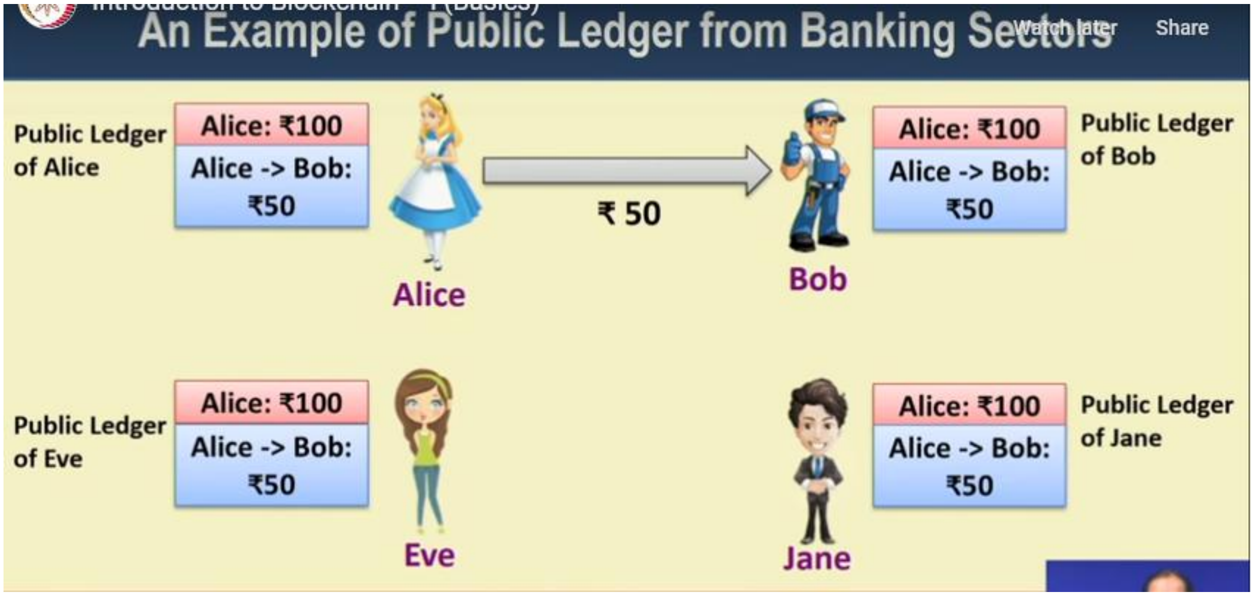


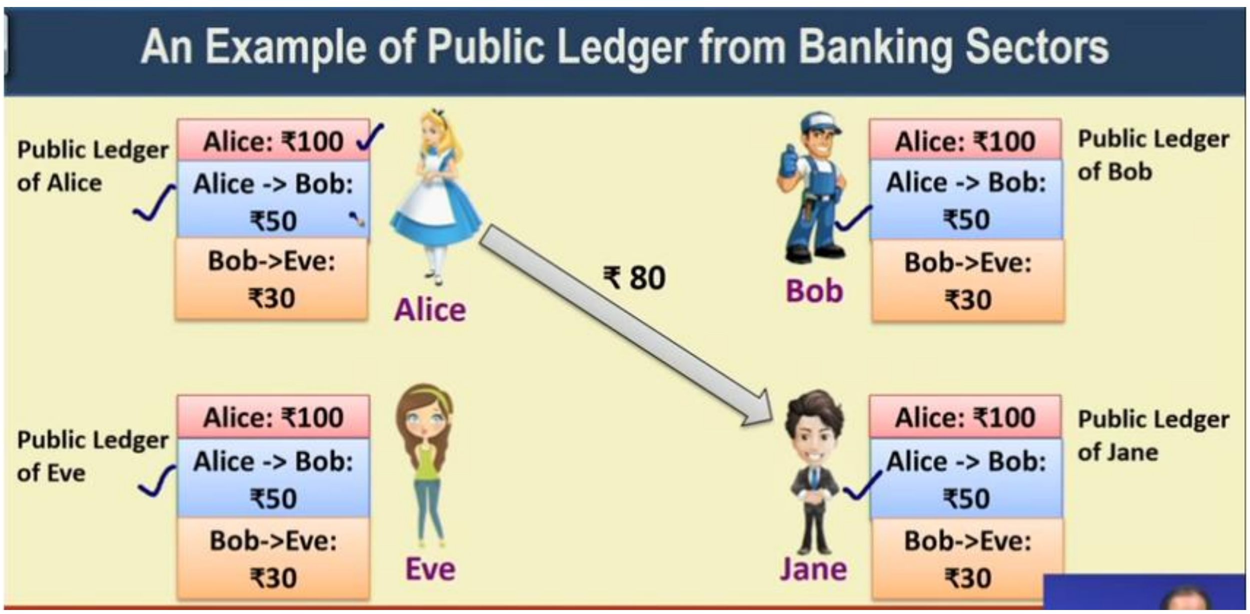
A Very Simplified Look of the Blockchain

- We call this a **Public Ledger**
  - A database of "**historical information**" available to everyone
  - The "**historical information**" may be utilized for future computation
- **An Example:**
  - Say, the historical information are the banking transactions
  - The old transactions are used to validate the new transactions

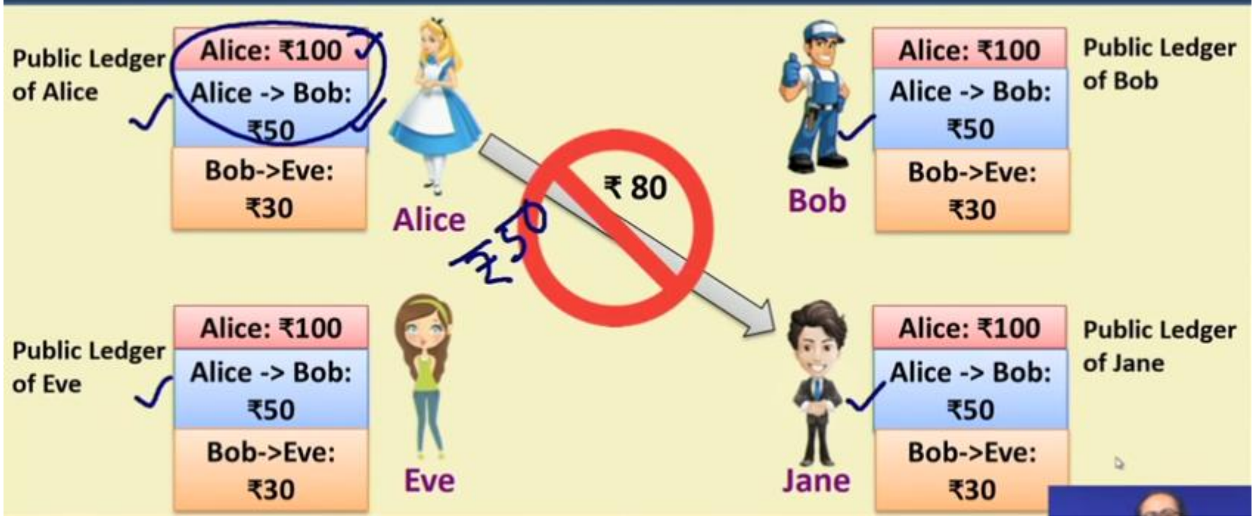


An Example of Public Ledger from Banking Sectors


An Example of Public Ledger from Banking Sectors
Watch later
Share
Public Ledger of Alice
Alice: ₹100
Alice -> Bob: ₹50
Alice
₹ 50
Bob
Alice: ₹100
Alice -> Bob: ₹50
Public Ledger of Bob
Public Ledger of Eve
Alice: ₹100
Alice -> Bob: ₹50
Eve
Jane
Alice: ₹100
Alice -> Bob: ₹50
Public Ledger of Jane


An Example of Public Ledger from Banking Sectors
Public Ledger of Alice
Alice: ₹100
Alice -> Bob: ₹50
Bob->Eve: ₹30
Alice
₹ 80
Bob
Alice: ₹100
Alice -> Bob: ₹50
Bob->Eve: ₹30
Public Ledger of Bob
Public Ledger of Eve
Alice: ₹100
Alice -> Bob: ₹50
Bob->Eve: ₹30
Eve
Jane
Alice: ₹100
Alice -> Bob: ₹50
Bob->Eve: ₹30
Public Ledger of Jane

An Example of Public Ledger from Banking Sectors



Blockchains and Public Ledgers

- Blockchains work like a public ledger
- However, we need to ensure a number of different aspects
  - **Protocols for Commitment:** Ensure that every *valid transaction* from the clients are committed and included in the blockchain within a finite time.
  - **Consensus**: Ensure that the local copies are consistent and updated.
  - **Security:** The data needs to be *tamper proof*. Note that the clients may act maliciously or can be compromised.
  - **Privacy and Authenticity:** The data (or transactions) belong to various clients; privacy and authenticity needs to be ensured.

Formal Definition of a Blockchain

- A Blockchain is "an **open**, **distributed ledger** that can record transactions between two parties **efficiently** and in a **verifiable** and **permanent** way" (Iansiti, Lakhani 2017)

- The keywords: **Open** (accessible to all), **Distributed or Decentralized** (no single party control), **efficient (**fast and scalable), **verifiable** (everyone can check the validity of information), **permanent** (the information is persistent)

Iansiti, Marco; Lakhani, Karim R. (January 2017). "The Truth About Blockchain". *Harvard Business Review*. Harvard University.

The Fundamentals

- **Cryptographically Secured Hash Functions**

  - ***Hash Functions***: Map any sized data to a fixed size; Example $H(x) = x \% n$, where x and n are integers and % is the modular (remainder after division by n) operations. x can be of any arbitrary length, but $H(x)$ is within the range $[0,n-1]$.

  - ***Cryptographically Secured:***
    - **One way**, given a x, we can compute $H(x)$, but given a $H(x)$, no deterministic algorithm can compute x
    - For two different x_1 *and* x_2, $H(x_1)$ and $H(x_2)$ should be different

Cryptographic Hash Functions

- Examples: MD5, SHA256
- X is called the **message** and H(X) is called the **message digest**
- A small change in the data results in a significant change in the output – called the **avalanche effect**

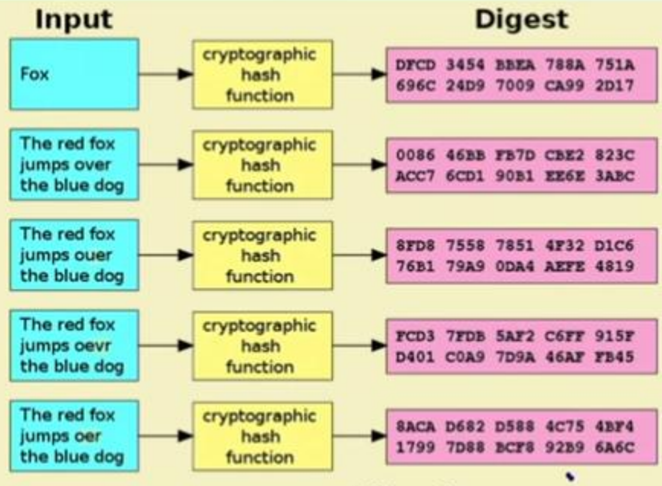
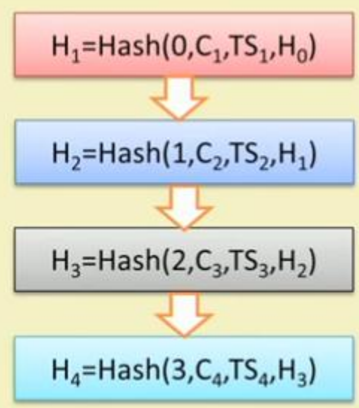


Image source: Wikipedia

Cryptographically Secured Chain of Blocks

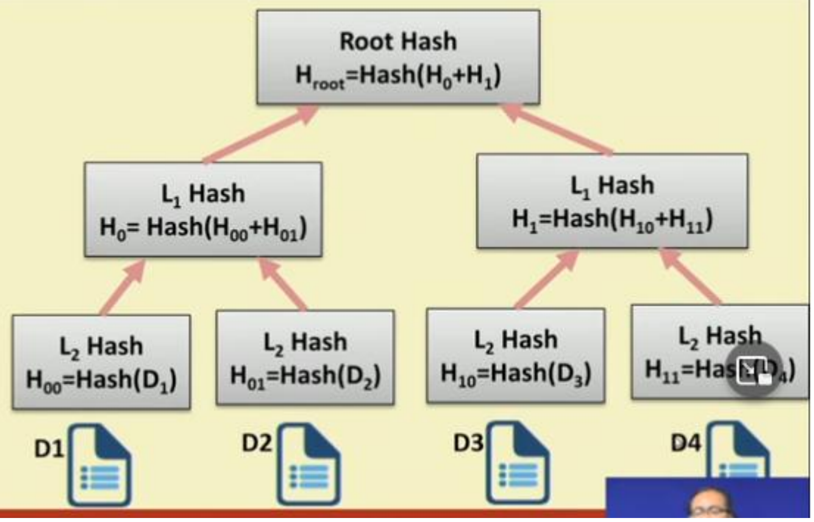
- The first use - **time-stamp a digital document** (*Harber and Stornetta, 1991*)
  - A sequence of timestamps [TS_1, TS_2, TS_3, …] denoting when the document is created or edited.
  - Whenever a client access a document, construct a block consisting of the sequence number of access, client ID, timestamp, a hash value from the previous request; and the entire thing is hashed to connect it to the previous blocks.



Haber, Stuart; Stornetta, W. Scott (January 1991). "How to time-stamp a digital document". *Journal of Cryptology*. 3 (2): 99–111

Merkle Trees (Ralph Merkle, 1979)

- Also known as **hash tree**
  - *every leaf node* is labelled with the hash of a data block
  - ***every non-leaf node*** is labelled with the cryptographic hash of the labels of its child nodes



Use of Merkle Trees

- Bayer, Harber and Stornetta used Merkle Tree in 1992 for timestamping and verifying a digital document - improved the efficiency by combining timestamping of several documents into one block

- Other uses of Merkle Tree
  - Peer to Peer Networks: Data blocks received in undamaged and unaltered; other peers do not lie about a block
  - **Bitcoin** implementation – shared information are unaltered; no one can lie about a transaction