Distributed hash tables

A Distributed Hash Table (DHT) plays a crucial role in enhancing the efficiency, scalability, and decentralisation of blockchain technology. Within blockchain networks, a DHT serves as a decentralised lookup system for key-value pairs, offering several key functionalities. Firstly, it facilitates peer discovery and message routing in peer-to-peer (P2P) networks, allowing nodes to efficiently find and communicate with each other without reliance on centralised servers or predefined peers. Additionally, DHTs enable decentralised storage solutions for blockchain data, particularly useful for off-chain data, smart contract states, and large files. By distributing data across the network, DHTs alleviate the burden on individual nodes, thus improving scalability. Moreover, these systems are well-suited for content addressing, ensuring data integrity and efficient retrieval based on cryptographic hashes. Projects like IPFS exemplify the integration of DHTs with blockchain technology, providing a decentralised, content-addressable method for storing and sharing data. In summary, the incorporation of Distributed Hash Tables empowers blockchain networks with robust peer discovery, efficient data storage, enhanced scalability, and resilient content addressing, all contributing to a more decentralised and efficient blockchain ecosystem.

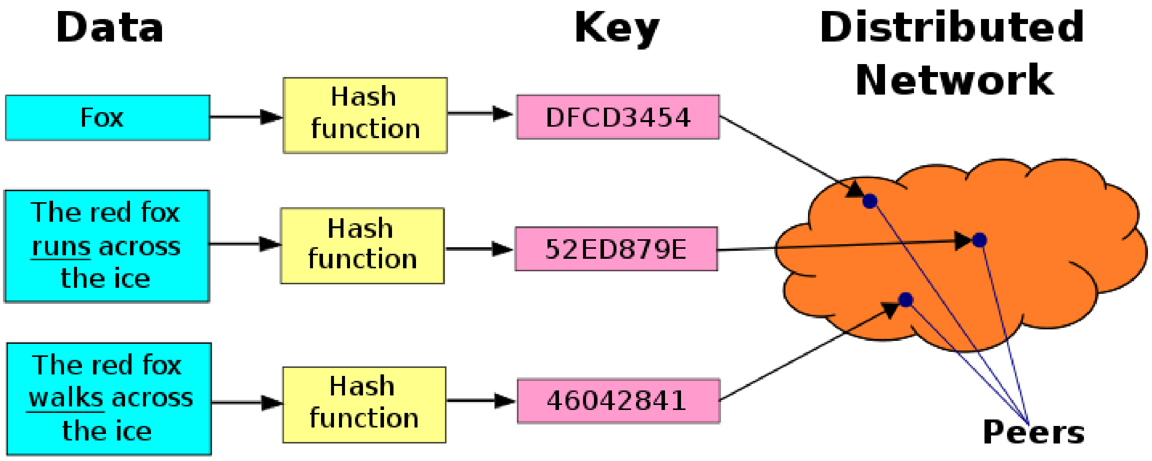

Fig 1.1. Example of a distributed hash table in a decentralised data store that holds data in key-value pairs.

Every key in a distributed hash table has a standard format and can be quickly divided into groups based on the location of each key and value pair, distributed hash tables offer a convenient approach to locate information among a vast collection of data. As each node stores the key partitioning scheme so that, in the event that it receives a request to access a certain key, it may promptly map the key to the node that stores the data, the nodes involved in a distributed hash table function as peers to find specific data values. After that, the request is sent to that node. Furthermore, adding or removing nodes from a distributed hash table is simple and doesn't need a lot of data rebalancing within the cluster. Cluster rebalancing is frequently a laborious process that affects performance, particularly for huge data sets. Applications that access data in the distributed hash table must not be hampered by changes in data size, which is why it is important to provide a fast and simple way to enlarge or shrink a cluster.

Blockchain and full ecosystem decentralisation

Blockchain technology and its full ecosystem are often associated with decentralisation, which is a key characteristic that distinguishes it from traditional centralised systems. Blockchain's core idea is the distributed ledger, which is not under the control of a single central authority but is instead kept up to date by a network of nodes. The blockchain ensures data immutability and transparency since every node in the network has a copy of the whole thing.

Peer-to-Peer Network:

Blockchains operate on a peer-to-peer (P2P) network where nodes communicate directly with each other without the need for intermediaries. This network architecture allows for data to be propagated and verified by multiple participants in a trustless manner. In a blockchain network, each participant, or node, holds a copy of the entire blockchain ledger, ensuring that no single entity controls the system. This distributed ledger is continuously updated as new transactions are added, and every node independently verifies and maintains the integrity of the blockchain.

Consensus Mechanisms:

Decentralised consensus mechanisms, such as Proof of Work (PoW), Proof of Stake (PoS), and others, enable nodes in the network to agree on the validity of transactions and the state of the ledger without relying on a central authority. Through these mechanisms, consensus is achieved among participants with potentially conflicting interests. These mechanisms ensure that the decentralised ledger remains consistent and secure, even in the absence of a central authority. Moreover, they contribute to the resilience and censorship resistance of blockchain networks, ensuring that they can operate autonomously without reliance on centralised authorities or intermediaries. As the blockchain ecosystem continues to evolve, consensus mechanisms will likely undergo further innovation and refinement to address scalability and security.

Storage:

Decentralisation is achieved because data can be directly stored in a blockchain. Nevertheless, a major drawback of this strategy is that a blockchain isn't meant to be used for storing massive volumes of data. It is undoubtedly unsuitable for storing images or big data blobs, as is the case with standard database systems. However, it can hold straightforward transactions and some arbitrary data. Distributed hash tables are a superior option for data storage (DHTs). Originally, peer-to-peer file sharing programs like BitTorrent, Napster, Kazaa, and Gnutella used DHTs. The CAN, Chord, Pastry, and Tapest1Y initiatives popularised DHT research. Although BitTorrent is the quickest and most scalable network, its disadvantage over other networks is that users are not incentivized to hold files for an extended period of time. Users generally don't keep files permanently, and if nodes that have data still required by someone leave the network, there is no way to retrieve it except by having the required nodes rejoin the network so that the files once again become available. High availability and link stability—that is, the ability to access network links at all times and the availability of data when needed—are the two main requirements here. Both of these qualities are present in Juan Benet's Inter- Planetary File System (IPFS), which aims to replace the HTTP protocol and create a decentralised World Wide Web. IPFS uses Merkle Directed Acyclic Graphs (DAGs) for searching and Kademlia DHT for storing. The Filecoin protocol serves as the foundation for the data storage incentive system. Filecoin offers rewards to nodes that store data via the Bitswap mechanism. Nodes can maintain a basic ledger of bytes delivered or received in a one-to-one relationship using the Bitswap technique. Additionally, IPFS uses a Git-based version management system to offer structure and control over data versioning. Other options for storing data include MaidSafe, Storj, and Ethereum Swarm. Ethereum has its own distributed and decentralised ecosystem that makes use of the Whisper protocol for communication and Swarm for storage. MaidSafe wants to offer a decentralised internet. BigChainDB is another storage layer decentralisation project aimed at providing a scalable, fast, and linearly scalable decentralised database as opposed to a traditional filesystem. BigChainDB complements decentralised processing platforms and filesystems such as Ethereum and IPFS.

Communication:

A decentralised system is the Internet, which is the communication layer of the blockchain. This idea was initially intended to create a decentralised communications network, so in that sense, it is partially accurate. Nowadays, the paradigm underlying services like email and online storage is one in which the service provider is in charge and consumers rely on them to provide access to the service when they need it. Since users do not have control over their data, this approach is predicated on the unwavering faith of a central authority (the service provider). It is on reliable third-party systems that user credentials are kept. Individual users must therefore be given power in a way that ensures their data access is secure and independent of a single third party. Internet Service Providers (ISPs), who serve as a central hub for Internet users, are the foundation for access to the Internet (the communication layer). With this paradigm, communication is impossible if the ISP is down for any reason. Alternatively, mesh networks can be used. They nonetheless offer a decentralised alternative to the Internet, where nodes can communicate directly with one another in the absence of a central hub like an ISP, despite having less capability than the latter. Globally, blockchain has resurrected the idea of decentralisation, and today, coordinated efforts are being made to utilise this technology and reap its advantages.

Computing power and decentralisation:

Blockchain technologies, like Ethereum, allow for the decentralisation of computing power by enabling smart contracts with embedded business logic to operate on the blockchain network. Similar processing-layer platforms are also offered by other blockchain systems, enabling decentralised business logic to be distributed throughout the network. This diagram provides a high-level representation of a decentralised ecosystem. A decentralised communication layer is offered by mesh networks or the Internet in the bottom layer. The storage layer at the top tier enables decentralisation with the help of BigChainDB and IPFS. At last, the subsequent stage reveals that the blockchain functions as a decentralised layer for processing, or computation. Blockchain can, in certain situations, offer a storage layer as well, but doing so significantly reduces the system's speed and capacity. For this reason, other approaches like BigChainDB and IPFS are more suited for decentralised data storage of massive volumes. At the top level, you can see the layers for Identity and Wealth. Internet identity is a broad topic, and systems:

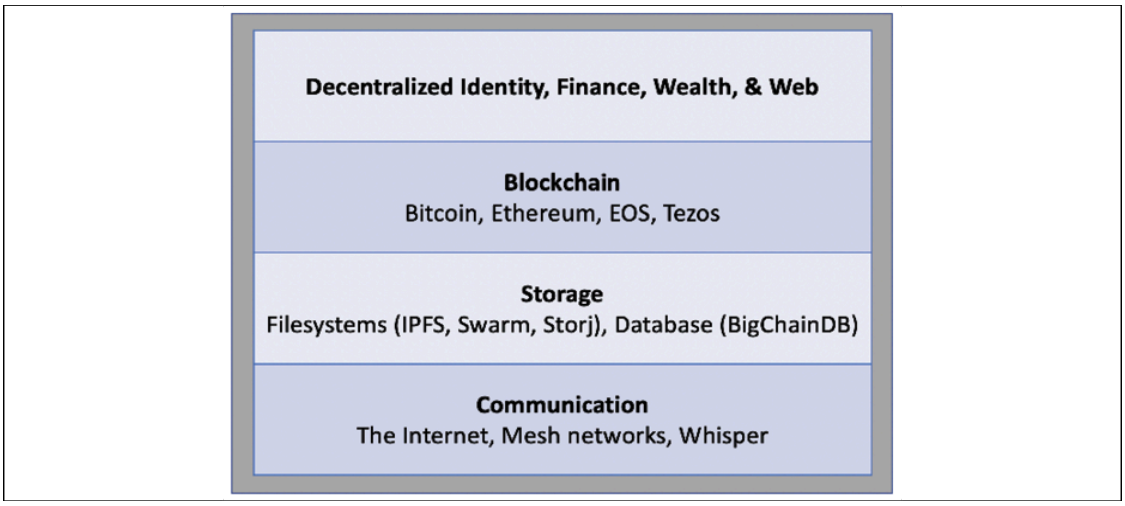
Decentralized Identity, Finance, Wealth, & Web
Blockchain
Bitcoin, Ethereum, EOS, Tezos
Storage
Filesystems (IPFS, Swarm, Storj), Database (BigChainDB)
Communication
The Internet, Mesh networks, Whisper

Fig 1.2. Decentralised Ecosystem

The blockchain has the ability to address a number of decentralisation-related problems. According to the identity-related Zooko's Triangle notion, a network protocol's naming system must be safe, decentralised, and capable of giving people names that have human meaning and are memorable. It is a conjecture that a system can only have two of these propellants operating at once. Nevertheless, with the advent of blockchain in the form of Namecoin, this problem was resolved. It is now possible to achieve security, decentralisation, and human-meaningful names with the Namecoin blockchain. However, this is not a panacea, and it comes with many challenges, such as reliance on users to store and maintain private keys securely. This opens up other general questions about the suitability of decentralisation to a particular problem. In many situations, decentralisation may not be the optimal choice. Established centralised systems, backed by reputable companies, often offer superior performance. For instance, email platforms provided by well-known companies like Google or Microsoft typically deliver better service compared to scenarios where individual users host their email servers on the Internet.