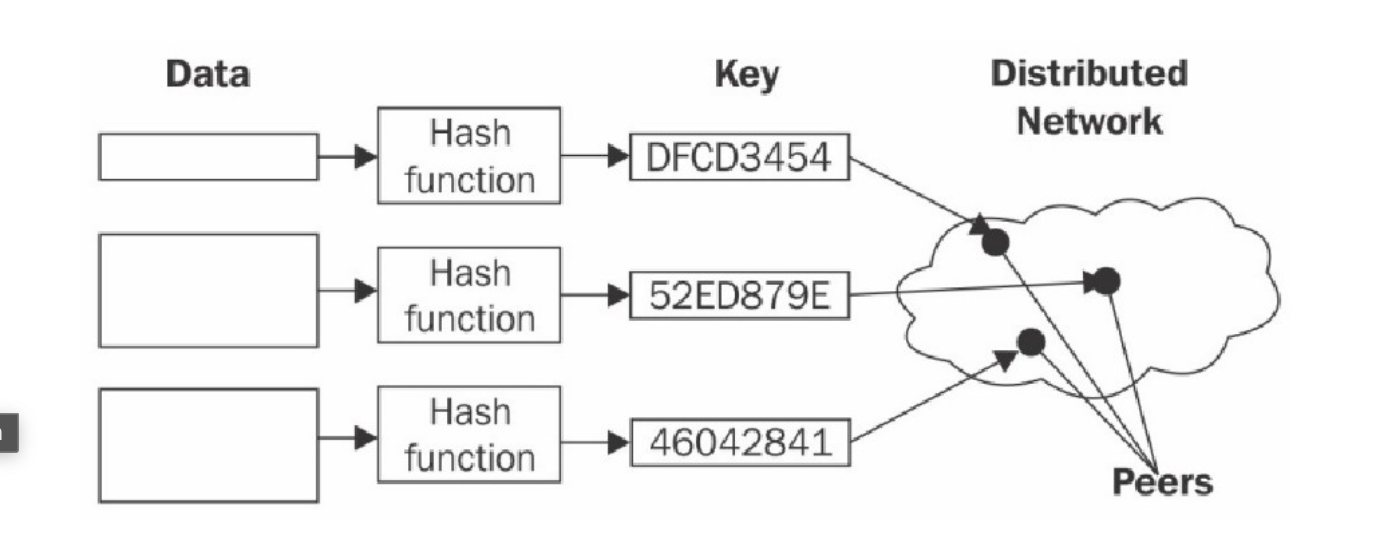
**Distributed hash tables**   
A distributed hash table (DHT) is a type of distributed system that provides a lookup service similar to a hash table. In a hash table, data is stored and retrieved using keys, and the keys are used to determine the location of the data in the table. A distributed hash table is similar, but the data is distributed across multiple nodes in a network rather than being stored in a single table.

In a DHT, each node is responsible for storing and managing a portion of the data. When a client wants to retrieve or store data, it sends a request to the network. The request is then forwarded to the appropriate node based on the key of the data being requested. The node then responds to the request and either retrieves or stores the data. DHTs are used in a variety of applications, including peer-to-peer (P2P) networks, distributed databases, and distributed file systems.

  
 Fig: Distributed Hash Table

**Blockchain and full ecosystem decentralization**

The blockchain is a distributed ledger that runs on top of conventional

systems. These elements include storage, communication, and computation.

There are other factors, such as identity and wealth, which are traditionally

based on centralized paradigms, and there's a need to decentralize these

aspects as well in order to achieve a sufficiently decentralized ecosystem.

**Storage**

Data can be stored directly in a blockchain, and with this fact it achieves

decentralization. However, a significant disadvantage of this approach is

that a blockchain is not suitable for storing large amounts of data by design.

It can store simple transactions and some arbitrary data, but it is certainly

not suitable for storing images or large blobs of data, as is the case with

traditional database systems.

A better alternative for storing data is to use distributed hash tables

(DHTs). DHTs were used initially in peer-to-peer file sharing software,

such as BitTorrent, Napster, Kazaa, and Gnutella. DHT research was made

popular by the CAN, Chord, Pastry, and Tapestry projects. BitTorrent is the

most scalable and fastest network, but the issue with BitTorrent and the

others is that there is no incentive for users to keep the files indefinitely.

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.

Two primary requirements here are high availability and link stability,

which means that data should be available when required and network links

also should always be accessible. Inter-Planetary File System (IPFS) by

Juan Benet possesses both of these properties, and its vision is to provide a

decentralized World Wide Web by replacing the HTTP protocol. IPFS uses

Kademlia DHT and Merkle Directed Acyclic Graphs (DAGs) to provide

storage and searching functionality, respectively. The concept of DHTs and

DAGs will be introduced in detail in Chapter 4, Public Key Cryptography.

The incentive mechanism for storing data is based on a protocol known as

Filecoin, which pays incentives to nodes that store data using the Bitswap

mechanism. The Bitswap mechanism lets nodes keep a simple ledger of

bytes sent or bytes received in a one-to-one relationship. Also, a Git-based

version control mechanism is used in IPFS to provide structure and control

over the versioning of data.

There are other alternatives for data storage, such as Ethereum Swarm,

Storj, and MaidSafe. Ethereum has its own decentralized and distributed

ecosystem that uses Swarm for storage and the Whisper protocol for

communication. MaidSafe aims to provide a decentralized World Wide

Web. All of these projects are discussed later in this book in greater detail.

BigChainDB is another storage layer decentralization project aimed at

providing a scalable, fast, and linearly scalable decentralized database as

opposed to a traditional filesystem. BigChainDB complements

decentralized processing platforms and filesystems such as Ethereum and

IPFS.

**Communication**

The Internet (the communication layer in blockchain) is considered to be

decentralized. This belief is correct to some extent, as the original vision of

the Internet was to develop a decentralized communications system.

Services such as email and online storage are now all based on a paradigm

where the service provider is in control, and users trust such providers to

grant them access to the service as requested. This model is based on the

unconditional trust of a central authority (the service provider) where users

are not in control of their data. Even user passwords are stored on trusted

third-party systems.

Thus, there is a need to provide control to individual users in such a way

that access to their data is guaranteed and is not dependent on a single third

party. Access to the Internet (the communication layer) is based on Internet

Service Providers (ISPs) who act as a central hub for Internet users. If the

ISP is shut down for any reason, then no communication is possible with

this model.

An alternative is to use mesh networks. Even though they are limited in

functionality when compared to the Internet, they still provide a

decentralized alternative where nodes can talk directly to each other without

a central hub such as an ISP.

Blockchain has revived the vision of decentralization across the world, and now concerted efforts are being made to harness this technology and take advantage of the benefits that it can provide.

**Computing power and decentralization**

Decentralization of computing or processing power is achieved by a

blockchain technology such as Ethereum, where smart contracts with

embedded business logic can run on the blockchain network. Other

blockchain technologies also provide similar processing-layer platforms,

where business logic can run over the network in a decentralized manner.

The following diagram shows an overview of a decentralized ecosystem. In

the bottom layer, the Internet or mesh networks provide a decentralized

communication layer. In the next layer up, a storage layer uses technologies

such as IPFS and BigChainDB to enable decentralization. Finally, in the

next level up, you can see that the blockchain serves as a decentralized

processing (computation) layer. Blockchain can, in a limited way, provide a

storage layer too, but that severely hampers the speed and capacity of the

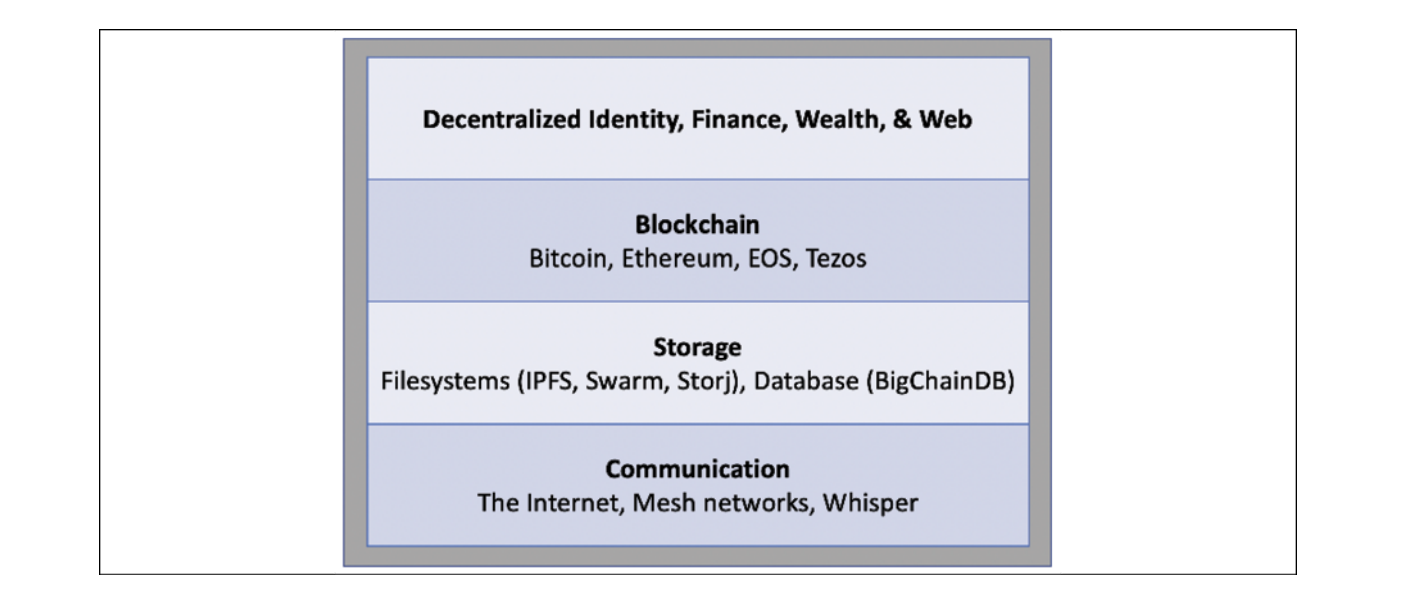
system. Therefore, other solutions such as IPFS and BigChainDB are more

suitable for storing large amounts of data in a decentralized way. The

Identity and Wealth layers are shown at the top level. Identity on the

Internet is a vast topic, and systems such as bitAuth and OpenID provide

authentication and identification services with varying degrees of decentralization and security assumptions

****

**Fig: Decentralized Ecosystem**

The blockchain is capable of providing solutions to various issues relating

to decentralization. A concept relevant to identity known as Zooko's

Triangle requires that the naming system in a network protocol is secure,

decentralized, and able to provide human-meaningful and memorable

names to the users. Conjecture has it that a system can have only two of

these properties simultaneously.

Nevertheless, with the advent of blockchain in the form of Namecoin, this

problem was resolved. It is now possible to achieve security,

decentralization, 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

decentralization to a particular problem.

Decentralization may not be appropriate for every scenario. Centralized

systems with well-established reputations tend to work better in many

cases.