

An Introduction to IPFS

Imagine a world where streaming a high-quality video on YouTube is bufferless, disrupted internet connection doesn't refrain you from working, where governments can't block any site or shut the entire internet down.

What are the issues with the current Web?

1. **Enormous Data** — With the advent of the internet, humanity has collected nearly 130 exabytes of data. This data is exponentially growing with the growing number of internet users. That means we will have more number of people streaming more data on channels that are not scaling at the same pace. This creates congestion which would not be solved by adding more hardware at one point of time.
2. **Centralized** — The current web is centralized which means that there is a single point of failure using which, the government of India shut down internet services in various parts of the country to damp down the fire of CAA and NRC protests.

Clearly, the problem exists. A solution to these problems is the use of a brand new internet protocol, IPFS.

IPFS

Juan Benet in 2014 designed an Internet Protocol commonly known as IPFS or Interplanetary File System with the objective of removing duplications across the network, keeping data indefinitely and acquire addresses to the information stored on computers in the network. The HTTP is location-oriented as it fetches information by pointing to locations. But IPFS is resource-oriented i.e. it points to the resource and fetches information in parts from different machines over the network making it decentralized.

So what difference does it make?

Suppose you are sitting in a classroom with 100 students and all of you are asked to stream the same 4k video on YouTube. The video will be retrieved from the closest YouTube server on your device and this process will be repeated a hundred times. Instead of sharing the video with each other, we transmit a large amount of data over long distances multiple times creating congestion and wastage of resources making HTTP inefficient.

Whereas IPFS tries to find data from the closest sources. This means that in the above scenario, the students would fetch the video from one another thereby making a 4k video stream bufferless.

But then how is IPFS different from the BitTorrent protocol?

BitTorrent has a separate group of users who share a single file or torrent. If a video exists in two different languages then there would be two swarms that could not share anything although both the swarms are carrying 99% content that is identical. But with the IPFS, the entire World Wide Web can be contemplated as a huge torrent file that is shared by everyone.

Also, BitTorrent does not support all data types and it contains duplicate data which adds up to the data congestion problems.

IPFS generates a unique cryptographic hash for every file submitted to it that depends on the content and hence achieves deduplication of data. It tracks version history that pre-empts information from being easily removed.

What are the misconceptions about IPFS?

1. **Data stored on IPFS is persistent** — Using cryptography, IPFS stores data on the network as long as the network believes that it is beneficial to do so. IPFS encourages users for data storage using cryptocurrencies like Filecoin.
2. **IPFS is built on Blockchain** — IPFS uses architectural elements similar to Merkle Trees but it is meant to work together with prevailing Blockchain protocols and not build on top of them. Blockchain provides publicly verifiability of data whereas IPFS provides publicly accessibility data.

Are there any applications with IPFS and Blockchain?

1. **Intellectual Property** — Arts, music, programs, source code etc. face problem of plagiarism. IPFS and blockchain can encourage content creators by providing a comprehensive ecosystem for them to work on. Revenue channels, in this case, can be driven by smart contracts, identity protection, a reputation-based collaboration network and more.
2. **Social Networks** — Blockchain and IPFS can provide users with an absolute decentralized social networking experience where content creation can be awarded with cryptocurrencies. Example, Akasha and Steemit.
3. **Free E-commerce** — Unlike traditional e-commerce sites like Amazon and e-Bay that charge fees for registering or selling items, an online marketplace can be created using IPFS that would not be centrally controlled and hence would not charge merchants. The trade would be peer-to-peer without any middleman. Example OpenBazaar.
4. **Cryptocurrencies** — Since a massive amount of data storage around the world sits unused, one could use it to store data for IPFS and earn FileCoin which can then be exchanged for other cryptocurrencies. These transactions are stored in blockchains.

How can IPFS beat the current web?

1. **Trustless multi-party computations** mean that computations can be performed on computers that aren't reliable. This applies to data storage as well. Doing this is not possible on today's web.
2. **Verifiable applications, timestamping and trusted ledgers** indicate that each operation that is happening can be mathematically proven.
3. **Decentralized applications** in the sense that an application on the internet can still exist without any infrastructure and support from the company even when the company cease to exist.
4. **End-to-end encryption is built-in.** Current providers subsist on user data and advertisement so it cannot offer end-to-end encryption whereas this is an in-built feature of web 3.0.