Week 5

Decentralized Storage

Cloud Storage

- Digital data storage itself has been around since the beginning of computers.
- From the 8KB data storage that put men on the moon, to the massive data center today; digital data storage is a critical component of today's economy and is one of the most lucrative businesses to be an early mover in.
- AWS, iCloud, Microsoft Azure, Google Cloud, and Dell EMC are some examples of cloud storage services.

Content Delivery Networks

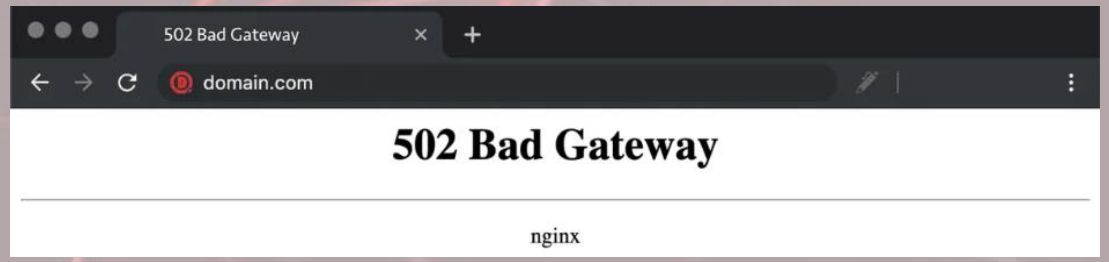
- When you watch a YouTube video, something on Netflix, or go to any website, you don't download the content that your computer displays.
- Data is delivered through complex networks of geographically distributed **data centers and proxy servers**. They started appearing in the 90s to fix the performance bottleneck of the internet.
- CDN providers use **physical cables** to deliver content at a faster rate than telephone lines used in the early days of the internet. They **do not host the content**, but rather provide a **superhighway** for the content to travel.
- **Cloudflare** is a good example of a CDN.
- Some companies like Google and Facebook have built their own CDN's with undersea cables to exclusively deliver their content.

Current storage meta

- Most people have either hear of or utilize “cloud storage.” Which is really just storage space held in massive datacenters by the respective companies holding the storage devices. The data **must** go **somewhere**.
- This includes AWS, Apple storage, OneDrive, and Snapchat memories (including My Eyes Only). The **data must go somewhere** if its not on your camera roll.☺
- These companies have total control over what happens with this data, and you are subject to the terms and conditions contracts **that you don’t read**.

Data Loss & Hacks

- Because cloud storage providers and content delivery networks are are often **centralized operations**, there is a **single point of failure** for many of these services, which makes them vulnerable to all types of attacks.
- Data centers in particular are vulnerable to **natural disaster** and destruction of the data facilities and cause huge data losses if it is not backed up somewhere safe.
- Content delivery networks are vulnerable to hacks and disruptions if the network is compromised. Recently, Cloudflare suffered an outage, which resulted in a lot of **502 errors** across the web.
- Data centers also represent a **gold mine** for hackers. If they are compromised, then untold amounts of data can be stolen.



What is decentralized storage?

Part 1

Permanent Records

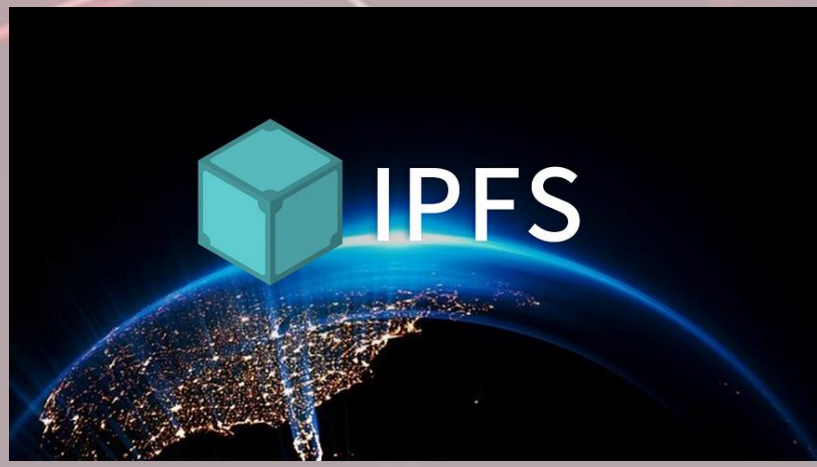
Secure, Encrypted Data

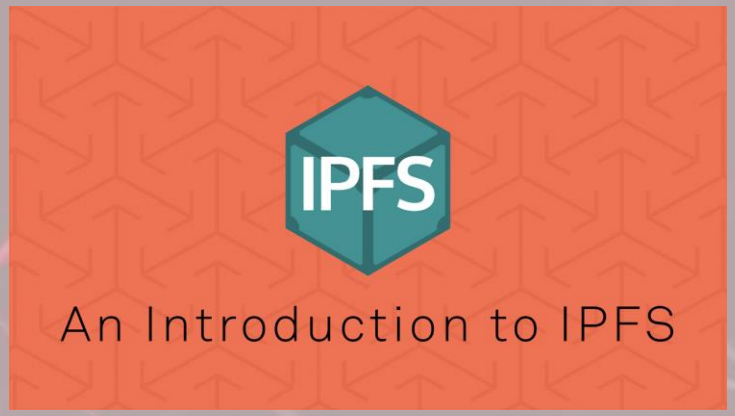
Business Model

- The business model for a decentralized storage protocol is one of the easier ones to comprehend in the world of crypto.
- They vary by platform, but the main thing to remember is that those who provide the **data storage space get rewarded**. This also applies to the nodes verifying transactions of the coin rewards, they get paid too for their computational power.

IPFS (Filecoin)

Part 2





What is IPFS?

- The **Interplanetary File System** is a distributed network designed for storing files, websites, applications, and data. Very similar to **BitTorrent**.
- A peer-to-peer hypermedia protocol, it is a filesharing network that uses **content-addressing** to uniquely identify each file in a global namespace connecting IPFS hosts.
- Instead of a traditional URL, it is something much different. Instead of asking Wikipedia's computers for the location of the Aardvark page, there is a mirror of it stored on IPFS.
- Also, instead of routing to a centralized server, your computer finds the closest IPFS node that looked up the same thing you are looking for.

```
https://en.wikipedia.org/wiki/Aardvark
```

```
/ipfs/bafybeiaysi4s6lnjev27ln5icwm6tueaw2vdykrtjkwiphwekaywqhcjze/wiki/Aardvark
```

Decentralization

- There are many benefits to making a decentralized network of storing content at multiple geographic points. It spreads out the risk and makes **multiple points of failure instead of one**.
- **Supports a resilient internet**. If someone attacks Wikipedia's web servers or an engineer at Wikipedia makes a big mistake that causes their servers to catch fire, you can still get the same webpages from somewhere else.
- **Makes it harder to censor content**. Because files on IPFS can come from many places, it's harder for anyone (whether they're states, corporations, or someone else) to block things. We hope IPFS can help provide ways to circumvent actions like these when they happen.
- **Can speed up the web when you're far away or disconnected**. If you can retrieve a file from someone nearby instead of hundreds or thousands of miles away, you can often get it faster. This is especially valuable if your community is networked locally but doesn't have a good connection to the wider internet.

Content Addressing

- The alphanumeric string of characters that comes after /ipfs/ is called a **content identifier**.
- **Location-based addressing** looks for the **location** of the file. For example; https://en.wikipedia.org/wiki/Aardvark. This method identifies a file by where it's located — what computer it's on and where on that computer's hard drive it is.
- IPFS uses something totally different. Instead of being location-based, IPFS addresses a file by **what's in it**, or by its **content**.
- Instead of asking **where** the content is, IPFS asks **what** the content is.
- The content identifier is the **cryptographic hash** of the content at that address. The hash is **unique** to the content that it came from, it is immutable and cannot be changed once its created.
- When you wish to change this content, such as in editing a website. It will generate a new alphanumeric string. This can be avoided with **DNS Link** and a **Mutable File System** if the content is going to change regularly. It also supports **versioning**.
- **IPLD (InterPlanetary Linked Data)** translates between hash-linked data structures, allowing for the unification of the data across distributed systems.

Filecoin

- Filecoin is the native token of IPFS and acts as the **incentive** for users to give the network their **raw storage space**. It is essentially a blockchain built on top of IPFS that creates a decentralized market for storage.
- If you have free space on your hard drive, you can rent it out for others to store files and make money in the process.
- Filecoin also makes sure that a piece of content is replicated on many nodes so there is a duplicate somewhere and the file is not lost.

Participation

- When you call a content address, the file will be called from the **closest** node running IPFS that called the content you are looking for.
- This makes the network very fast, as it does not need to travel back and forth between your computer and a server on the other side of the globe. It saves up to **60%** of bandwidth.
- All content published on IPFS is public by default, although this can be changed by encrypting a file, only allowing a select few people to see it.
- To participate in the network, you can either run as an IPFS node granting you full access to the network, pay a node operator to store your data, or use APIs to communicate with IPFS.
- Currently, only **Brave Browser** can understand /ipfs/ URLs; other browsers must include **https://ipfs.io/** before the IPFS hash.

Directed Acyclic Graphs (Merkle DAGs)

- Many distributed systems take advantage of **data structures** called DAGs.
- IPFS uses a Merkle DAG that is optimized for representing directories and files, but you can structure a Merkle DAG in many ways. For example, **Git** uses a Merkle DAG that has many versions of your repo inside of it.

Distributed Hash Tables (DHTs)

Libp2p

- What makes libp2p especially useful for peer-to-peer connections is **connection multiplexing**.

Use Cases

- In 2017, the Turkish government blocked Wikipedia because they deemed it a “threat to national security.” The Turkish Wiki was then promptly uploaded to IPFS where is could not be censored.
- DTube is very close to a YouTube clone that runs on IPFS. It is completely uncensorable by design.

0Chain

Part 3

What is 0Chain? (ZCN)

dStorage

- The 0Chain protocol writes data to **dStorage** in parallel chunks of 64kB and are cryptographically signed by the originator as **write markers**.
- These markers are then rolled up to a Merkle root of the **file** then to a Merkle root of the **allocation**.
- Blobbers then submit these markers to the blockchain for all operations of create/writes, reads, updates and deletes (known as CRUD operations) in order to submit the final state of the allocation.

CRUD Operations

Blobbers

Miners

Sharders

ZCN Token

Zbox CLI

Zwallet CLI

Other Projects

Part 4

BitTorrent

Sia

Storj