Universidad Politécnica
de Madrid

Escuela Técnica Superior de Ingenieros Informáticos

Grado en Grado en Ingeniería Informática

Trabajo Fin de Grado

Desarrollo de un Sistema de Intercambio Directo de Archivos entre Dispositivos Basado en IPFS

Autor: Nicolás Cossío Miravalles
Tutor(a): Fernando Pérez Costoya

Madrid, Abril - 2023

Este Trabajo Fin de Grado se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Grado
Grado en Grado en Ingeniería Informática

Título: Desarrollo de un Sistema de Intercambio Directo de Archivos entre Dispositivos Basado en IPFS

Abril - 2023

Autor: Nicolás Cossío Miravalles
Tutor: Fernando Pérez Costoya
Arquitectura Y Tecnología De Sistemas Informáticos
ETSI Informáticos
Universidad Politécnica de Madrid

Resumen

IPFS, también conocido como Protocolo de Sistema de Archivos Interplanetario, es un protocolo de red y un sistema de archivos diseñado para hacer la web más rápida, segura y abierta. Este sistema permite a los usuarios no solo recibir, sino también alojar contenido en una red P2P completamente descentralizada.

IPFS tiene varias ventajas clave. A diferencia de protocolos como HTTP, en IPFS los archivos se identifican por su contenido en lugar de por su ubicación. Esta característica permite a cualquier nodo de la red convertirse en proveedor de contenido dentro de ella, lo que se traduce en una mayor eficiencia, seguridad, escalabilidad y resiliencia para el almacenamiento y distribución de datos. IPFS facilita la creación de aplicaciones descentralizadas (dApps) al proporcionar herramientas como un sistema de almacenamiento de archivos distribuido y un sistema de nombres descentralizado (IPNS) para la web. Al mismo tiempo, promueve el desarrollo de aplicaciones resistentes a la censura y una web verdaderamente abierta y descentralizada.

Este trabajo de fin de grado se divide en dos partes:

La primera consiste en el estudio del ecosistema de IPFS. Se abarca desde su arquitectura, algoritmo de intercambio de bloques, identificación basada en contenido, hasta su estructura de datos. Se analizan ejemplos de casos de uso en la Web3, como la distribución descentralizada de contenido, el almacenamiento de datos en la cadena de bloques y la publicación de datos permanentes.

La segunda parte del trabajo consiste en la creación de un sistema de intercambio de archivos basado en IPFS. Se presenta un posible diseño de una arquitectura centralizada típica que se usaría para una aplicación de intercambio seguro de archivos. Se profundiza en posibles puntos únicos de falla, preocupaciones de privacidad y problemas de escalabilidad que surgen al depender de una sola autoridad o servidor.
Con estos puntos establecidos, se introduce el sistema ideado. Empleando la naturaleza distribuida de IPFS, esta propuesta tiene como objetivo abordar los problemas mencionados, minimizando los puntos únicos de falla y mejorando la propiedad y la privacidad de los datos.

Algunas características fundamentales de este sistema son:

- Archivado y compresión de archivos y directorios utilizando tar y gzip.
- Encriptación segura de archivos utilizando aes-256-cbc.
- Encriptación de secretos facilitada por JSON Web Encryption (JWE).
- Verificación de autoría mediante el uso de Identificadores Descentralizados (DIDs), en forma de firma de contenido utilizando JSON Web Signatures (JWS).
- Uso de bases de datos descentralizadas impulsadas por OrbitDB, que permiten:

- Silos de usuarios, registro automático y controladores de acceso distribuidos.
- Notificaciones push mediante una cola de mensajes descentralizada.
- Bases de datos locales con persistencia para uso interno de la aplicación.

La aplicación desarrollada funciona en sistemas operativos Windows, MacOS y Linux. Mediante una interfaz de comandos de consola los usuarios pueden compartir archivos de manera segura y privada sin la necesidad de depender de servidores centralizados.

Abstract

IPFS, also known as the InterPlanetary File System, is a network protocol and file system designed to make the web faster, more secure and open. This system allows users not only to receive but also to host content on a fully decentralized peer-to-peer network.

IPFS has several key advantages. Unlike protocols like HTTP, in IPFS, files are identified by their content rather than their location. This feature allows any node in the network to become a content provider, resulting in greater efficiency, security, scalability and resilience for data storage and distribution.

IPFS facilitates the creation of decentralized applications (dApps) by providing tools such as a distributed file storage system and a decentralized naming system (IPNS) for the web. At the same time it promotes the development of censorship-resistant applications as well as a truly open and decentralized web.

This undergraduate thesis is divided into two parts:

The first part consists of the study of the IPFS ecosystem. From its architecture, block exchange algorithm, content-based addressing, to its data structure. Examples of use cases in Web3, such as decentralized content distribution, blockchain-based data storage, and permanent data publishing, are also analyzed.

The second part of the thesis involves the creation of a secure and decentralized file-sharing system based on IPFS. The process starts by outlining the design and limitations of a typical centralized architecture for an application of the proposed type. Emphasizing on potential single points of failure, privacy concerns and scalability issues that arise from relying on a single authority or server.
With these points established, the devised system is then introduced. Employing the distributed nature of IPFS, this proposal aims to address the aforementioned issues, minimizing single points of failure, and enhancing data privacy and ownership.

Fundamental features of this system encompass:

- File or directory archiving and compression using tar and gzip.
- Secure file encryption using aes-256-cbc.
- Secrets encryption facilitated by JSON Web Encryption (JWE).
- Authorship verification through the usage of Decentralized Identifiers (DIDs), in the form of content signing using JSON Web Signatures (JWS).
- Usage of decentralized databases powered by OrbitDB which enable:
  - User silos, automatic registration, and distributed access controllers.
  - Push notifications via a decentralized message queue.
  - Local databases with persistence for internal application use.

The developed application supports Windows, MacOS, and Linux operating systems. Through a command-line interface, users can securely and privately share files without relying on centralized servers.

Tabla de contenidos

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

El presente Trabajo de Fin de Grado (TFG) se centra en el desarrollo de un sistema de intercambio de ficheros basado en IPFS (InterPlanetary File System)[**?**].
A continuación, se describen las motivaciones y necesidades que han llevado a la realización de este proyecto.

1.1. Motivación y necesidad

El desarrollo de un sistema de intercambio de ficheros basado en IPFS se encuentra en la confluencia de varias tendencias tecnológicas y sociales que están dando forma al futuro de la web. En particular, este proyecto se relaciona estrechamente con el avance hacia la *Web3*[**?**], una visión de un internet más descentralizado, seguro y resistente a la censura. En esta sección, exploraremos cómo un sistema de intercambio de archivos encaja en este nuevo panorama y por qué es relevante para el progreso de la Web3.

Los servicios de almacenamiento y compartición de archivos actuales, como Google Drive, Dropbox, Microsoft OneDrive y otros proveedores de almacenamiento en la nube son servicios centralizados y, aunque populares y ampliamente utilizados debido a su facilidad de uso, accesibilidad y confiabilidad, presentan ciertos problemas y limitaciones. Los usuarios dependen de una sola entidad para almacenar y gestionar sus archivos, lo que puede generar problemas si la empresa experimenta fallos técnicos, cambia sus políticas de uso, o se convierte en el objetivo de un ataque cibernéticos maliciosos. Además, esto otorga a estas empresas un gran poder sobre los datos de los usuarios, lo que puede conducir a problemas de privacidad y control de la información.

Otras alternativas como FTP (File Transfer Protocol) ofrecen una mayor autonomía y control sobre los archivos, pero también tienen inconvenientes. FTP es un protocolo que permite la transferencia de archivos entre un cliente y un servidor a través de una red. FTP carece de robustas medidas de seguridad modernas, puede ser vulnerable a ataques y requiere un mayor conocimiento técnico y esfuerzo para su configuración y mantenimiento.
En resumen, a pesar de la mayor autonomía y control directo que FTP puede ofrecer, no es comparable con un servicio en la nube en términos de seguridad, facilidad de uso y eficiencia de costos. Esto es teniendo en cuenta los conocimientos y requisitos

del usuario promedio de un servicio de estas características.

La arquitectura detrás de este tipo de servicios se basa en el modelo cliente-servidor. En este modelo, un servidor central almacena la información sobre la lista de nodos y recursos disponibles en la red y es vital para el funcionamiento del sistema. Esto facilita encontrar rápidamente los nodos o recursos disponibles, pero el sistema es relativamente vulnerable en términos de fallos o ataques y la escalabilidad está limitada debido a la presión sobre el elemento central [**?**].

La alternativa a estos servicios centralizados es el uso de tecnologías *peer-to-peer* (de igual a igual en español). Una aplicación peer-to-peer (p22) es un tipo de red donde no existen clientes ni servidores fijos, sino una serie de nodos que actúan como iguales y pueden funcionar tanto como clientes como servidores entre sí.

Existen varias tecnologías p2p que permiten compartir archivos entre usuarios sin necesidad de un proveedor central como los previamente mencionados, el más famoso y conocido siendo BitTorrent[**?**]. Sin embargo, estas tecnologías no son adecuadas para el intercambio de archivos entre usuarios no conocidos, ya que requieren que los usuarios confíen en que los archivos que se comparten son los que se anuncian.

Esto es algo que resuelve el Inter Planetary File System (IPFS). El Sistema de Archivos Interplanetario es un sistema de archivos distribuido que busca conectar todos los dispositivos al mismo sistema de archivos. En cierto modo, IPFS es similar a la Web, aunque podría verse como una sola red BitTorrent, intercambiando objetos dentro de un repositorio Git.
En otras palabras, IPFS permite guardar y acceder a bloques de datos identificados por su contenido, no por su ubicación, y que se pueden transferir rápidamente entre los nodos. Además, IPFS usa estos bloques para crear enlaces que también se basan en el contenido, no en una dirección que apunta a una ubicación donde se puede encontrar el contenido. Esto forma un grafo dirigido acíclico generalizado de Merkle (Merkle DAG), una estructura de datos sobre la que se puede construir sistemas de archivos versionados, cadenas de bloques e incluso una Web Permanente. IPFS combina una tabla hash distribuida, intercambio de bloques incentivado y espacio de nombres autocertificante, sin puntos únicos de falla ni necesidad de confianza entre los nodos que la forman[**?**].

En este proyecto se usará IPFS como bloque central, sobre el que construirá el sistema previamente descrito.

1.2. Objetivos y alcance del proyecto

El objetivo principal de este proyecto es el desarrollo de un sistema de intercambio de ficheros basado en IPFS, mediante el desarrollo una aplicación de escritorio. Este sistema debe permitir a los usuarios compartir archivos de forma segura y confiable, sin necesidad de ningún proveedor central de ningún tipo.
Debe integrar capacidades de encriptación y control de acceso para garantizar la seguridad de los archivos compartidos. La integración de cuentas de usuario, con la posibilidad de hacer grupos, elegir contactos con los que compartir, se propone como algo imprescindible para lograr un sistema autocontenido y sin necesidad de herramientas externas para su uso. Por último se debe integrar un sistema de notificaciones para el que los usuarios puedan recibir avisos de nuevos archivos compartidos,

o de cambios en los archivos compartidos.

Para lograr esto se han cumplido los siguientes objetivos:

- Investigar sobre IPFS y su funcionamiento para entender cómo funciona el protocolo y cómo se puede utilizar para el sistema propuesto.
- Investigar sobre el ecosistema en torno a IPFS, con objetivo de comprender la madurez y viabilidad de esta tecnología, así como de las herramientas basadas en esta que se pueden utilizar para el sistema propuesto.
- Diseñar una arquitectura para el sistema de intercambio en torno a las tecnologías y herramientas seleccionadas.
- Implementación de un prototipo funcional del sistema propuesto.
- Analizar la viabilidad de IPFS en base a la experiencia obtenida en el desarrollo del prototipo.
- Analizar posibles mejoras y ampliaciones del sistema propuesto.

Por tanto pese a que el objetivo principal es el desarrollo de un sistema de intercambio de ficheros basado en IPFS, también se realizará una labor de investigación sobre IPFS y su ecosistema, con el objetivo de comprender esta tecnología y su viabilidad como alternativa a muchas de las tecnologías actuales.

Sobre el alcance del proyecto, en el capitulo **??**: '**??**' se explora las posibles vías de expansión y mejoras para el proyecto en el futuro. También se expresan las esperanzas y expectativas para el crecimiento y posible impacto del mismo.

1.3. Estructura de la memoria

En este capítulo se ha introducido el proyecto, explicando las motivaciones y necesidades que han llevado a su realización.

En el capítulo **??**: '**??**' se pone en situación el estado actual de tecnologías relacionadas con el proyecto, tanto alternativas como otras implementaciones que usen IPFS u otras tecnologías similares que cumplan parcial o completamente con los objetivos del proyecto. Al comienzo de este capítulo también se explica brevemente la historia de internet y su evolución hasta el presente. La razón de ser de esta sección se debe a la necesidad de poner en situación el porqué detrás de la dominancia de ciertos protocolos que han guiado el modelo de internet actual, y que han llevado a la necesidad de alternativas como IPFS.

Dentro de este capítulo se explica el funcionamiento de IPFS, tratando los siguientes temas: arquitectura interna, funcionamiento, ecosistema y herramientas relacionadas. Con esta sección se busca dar una visión general de esta tecnología y su ecosistema para poder entender el sistema propuesto.

El capítulo **??**: '**??**' se centra en el desarrollo del sistema propuesto. Este se ha estructurado en en:

- **Requisitos del sistema:** se explica el funcionamiento deseado del sistema.
- **Diseño del sistema:** se presenta la arquitectura y diseño propuestos en este proyecto, así como las herramientas utilizadas.

- **Implementación:** se explica la implementación realizada, así como las decisiones tomadas durante el desarrollo. Esta sección incluye partes de código relevantes para entender la implementación realizada.

En el capítulo **??**: '**??**' se realiza una serie de pruebas del sistema desarrollado. Para ello se ha creado un escenario de uso real con distintos usuarios en varios lugares del mundo.

El capítulo **??**: '**??**' se analiza el resultado obtenido del desarrollo del proyecto. Se contrastará el resultado con los objetivos propuestos y con servicios de transferencia de archivos centralizados.

En el capítulo **??**: '**??**' se proponen posibles mejoras y ampliaciones del sistema propuesto.

Capítulo 2

Contexto y estado del arte

En esta sección se intentarán poner en perspectiva, de una forma no exhaustiva, las distintas razones históricas que dan lugar a la necesidad de crear un sistema de almacenamiento descentralizado y distribuido como IPFS. Para ello, se hará un breve repaso histórico de la evolución de Internet y de los protocolos que lo han ido conformando. Posteriormente, se explicará la tendencia centralista del sistema actual, existente a un nivel intrínseco y estructural, además de otros problemas que se derivan de esta situación. Finalmente, se expondrá la propuesta de solución que IPFS ofrece para solventar estos problemas, sobre la que se profundizará en la sección **??**: '**??**'.

2.1. Breve historia de Internet

2.1.1. Predominancia de los protocolos TCP/IP

La historia de internet está marcada por la competencia entre distintos protocolos de comunicación que buscaban establecerse como el estándar para intercambiar información entre diferentes redes y sistemas. Uno de los episodios más relevantes de esta competencia fue la llamada *"Guerra de los protocolos"* [**?**], en la que el conjunto de protocolos TCP/IP, creado entre los años 1973 y 1974 por Vint Cerf y Robert Kah, se enfrentó a otras propuestas como OSI, X.25 o SNA[1].

TCP/IP logró imponerse a la competencia debido a las siguientes características principalmente:

- **Interoperabilidad** : La capacidad de TCP/IP para conectarse fácilmente con diferentes tipos de ordenadores y sistemas operativos le otorgaba una ventaja sobre otros protocolos que eran más específicos o limitados en su compatibilidad. Esta característica permitía que diversas tecnologías y plataformas pudieran comunicarse entre sí sin problemas, lo cual era esencial para crear una red global como internet.
- **Flexibilidad** : TCP/IP podía adaptarse a distintos medios de transmisión, como cables de cobre, fibra óptica o incluso enlaces inalámbricos, lo que facilitaba

[1] En la figura **??** de la página **??** se muestra un resumen de las principales características de cada uno de estos protocolos.

su implementación en una amplia variedad de entornos y situaciones. Otros protocolos, en cambio, podrían haber requerido modificaciones o adaptaciones específicas para funcionar en diferentes tipos de medios de transmisión.

- **Resistencia** frente a fallos: TCP/IP fue diseñado para ser robusto en caso de fallos en la red, permitiendo que los paquetes de datos pudieran ser retransmitidos y encontrar rutas alternativas en caso de problemas. Esta capacidad de recuperación era fundamental para garantizar la continuidad y fiabilidad de las comunicaciones en una red global con múltiples nodos y enlaces.

- **Escalabilidad** : TCP/IP podía soportar el crecimiento de la red al permitir la incorporación de nuevos nodos y enlaces sin afectar negativamente su rendimiento. Su diseño jerárquico y descentralizado facilitaba la expansión de la red y evitaba los cuellos de botella que podrían haberse producido con otros protocolos menos escalables.

Estas ventajas hicieron que TCP/IP se convirtiera en la opción preferida frente a otros protocolos, al ser una solución más versátil, resistente y escalable para la creciente demanda de interconexión entre sistemas y redes en todo el mundo. Cabe destacar también que era una solución con arquitectura abierta, no propietaria y de uso gratuito, es decir, sin necesidad de pagar licencias por su uso [**?**].

Como en toda guerra también hubo un trasfondo político. Este hecho suele ser ignorado al abordarse este tema desde un punto de vista puramente tecnológico. Y es que en 1980, el Departamento de Defensa de Estados Unidos declaró TCP/IP como el estándar para todas las redes militares [**?**]. A esto se sumaron numerosas comunidades de investigación y universidades que adoptaron TCP/IP como su protocolo de comunicación, como por ejemplo, Stanford University, donde Vint Cerf colaboró con Robert Kahn en el diseño del protocolo [**?**]; University of California, Los Angeles (UCLA), que participó en el desarrollo temprano y las pruebas de TCP/IP [**?**]; y University College London (UCL), donde el profesor Peter Kirstein promovió el uso de TCP/IP en Europa y su equipo contribuyó al desarrollo y pruebas del protocolo [**?**].

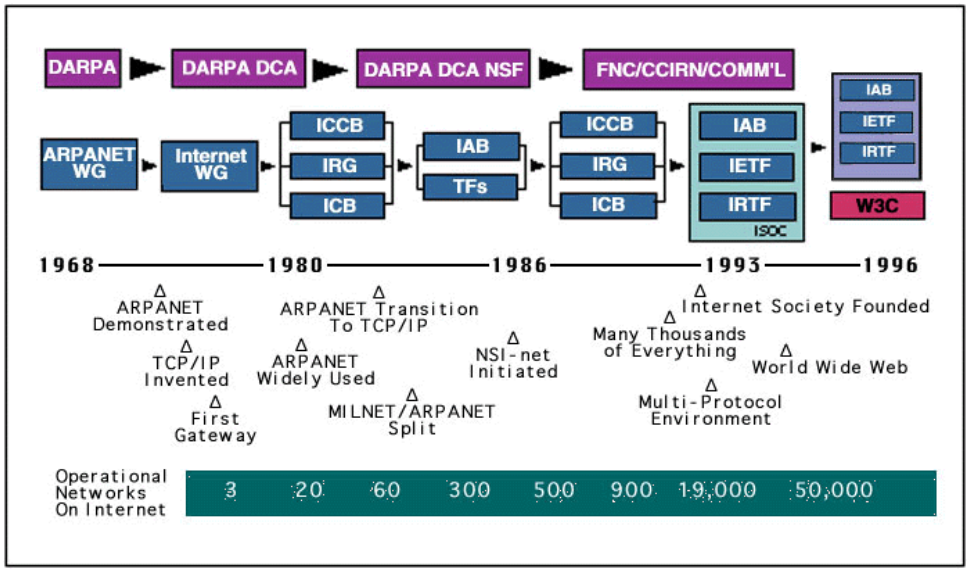

Figura 2.1: Evolución de los protocolos de Internet. Fuente [**?**]

Esta completa adopción del protocolo se dió por finalizada cuando ARPANET precursor de internet y financiado por la Agencia de Proyectos de Investigación Avanzados de Defensa (DARPA), llevó a cabo la transición exitosa de su antiguo protocolo, el Network Control Program (NCP), a TCP/IP el 1 de enero de 1983 [**?**].

En resumen, la rápida adopción de la comunidad científica y académica, sumada al respaldo gubernamental consolidaron TCP/IP como el estándar dominante en la industria de las redes de comunicación.

2.1.2. La World Wide Web y HTTP

El modelo TCP/IP asentó una forma de comunicación estándar entre computadores y redes, aunque este estaba limitado principalmente al mundo académico y científico. No fue hasta la creación de la World Wide Web (WWW) cuando el Internet concebido como es en la actualidad se convirtió en un fenómeno global y accesible para todo el mundo.

Antes de la WWW, el acceso a la información en Internet se realizaba a través de los protocolos a nivel de aplicación mostrados en la figura **??**

Protocolo	**Descripción**
FTP (Protocolo de Transferencia de Archivos)	Utilizado para transferir archivos entre cliente y servidor a través de una red.
Telnet	Basado en texto utilizado para el acceso remoto a computadoras y servidores, permitiendo a los usuarios controlarlos a través de una interfaz de línea de comandos.
Gopher	Diseñado para buscar y recuperar documentos de manera jerárquica, utilizando una interfaz basada en menús.
SMTP (Protocolo Simple de Transferencia de Correo)	Utilizado para enviar mensajes de correo electrónico entre servidores y, finalmente, al cliente de correo del destinatario.
NNTP (Protocolo de Transferencia de Noticias en Red)	Utilizado para la distribución, consulta y recuperación de artículos de noticias en la red Usenet.
POP3 (Protocolo de Oficina de Correos 3)	Utilizado para recuperar mensajes de correo electrónico desde un servidor de correo remoto hasta un cliente de correo local.
IMAP (Protocolo de Acceso a Mensajes de Internet)	Permite a los usuarios acceder y administrar sus mensajes de correo electrónico en un servidor de correo, sin descargarlos a un cliente de correo local.

Cuadro 2.1: Protocolos de capa de aplicación antes de HTTP

Estos servicios se encuentran en nivel de aplicación dentro del *stack* TCP/IP, como se muestra en la figura **??**. Algunos de ellos se siguen usando hoy en día, o tienen su caso de uso (IMAP, POP3, FTP), pero en lo referente a archivos, ofrecían métodos básicos de navegación y compartición. Carecían de la capacidad de inter-conectar

documentos de manera intuitiva y visual. [2]

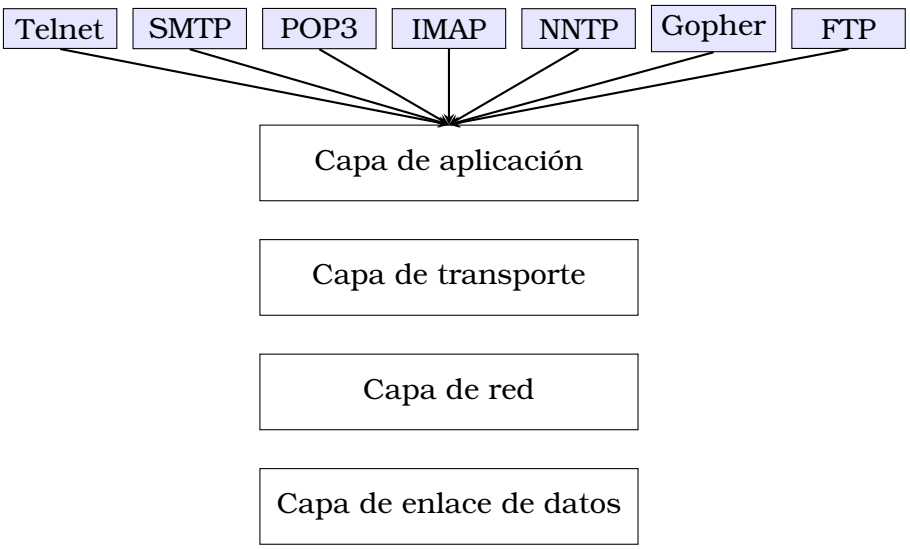

Figura 2.2: Capas del protocolo TCP/IP mostrando algunos protocolos de la capa de aplicación

En 1989, el científico británico Tim Berners-Lee propuso la creación de la WWW, un sistema de información global que permitiría a los usuarios navegar y acceder a documentos interconectados mediante enlaces. Estos documentos, conocidos como páginas web, se almacenarían en computadoras conectadas a la red y podrían ser accedidos a través de un programa especial llamado navegador web, que interpretaría el código de las páginas y mostraría su contenido al usuario.
HTML (Hyper Text Markup Language) es el lenguaje que describe estos documentos. Permite enlazar documentos entre sí mediante hipervínculos. Un hipervínculo es una referencia unidireccional en un documento electrónico que entrelaza diferentes documentos o secciones entre sí. Los usuarios tienen la oportunidad de seguir estos enlaces con tan solo un clic en el texto ancla (texto enlazado) para navegar a los documentos o las secciones correspondientes[**?**]. Aunque es un concepto simple y con el que cualquier persona en la actualidad esta familiarizada este factor dictamina la forma en la que se usa internet en la actualidad. Los usuarios de internet interactúan con el contenido en internet mediante estos enlaces.
La WWW se basó en tres tecnologías clave: HTML, un lenguaje de marcado para crear páginas web; HTTP, un protocolo para solicitar y transferir recursos a través de la web; y URL, un sistema de direcciones para localizar recursos en la web[**?**]. Y es este último el que genera una gran problemática que resuelve IPFS.

URL significa Uniform Resource Locator, que se traduce al español como Localizador Uniforme de Recursos. Es un sistema de direcciones utilizado en la web para localizar de manera única recursos como páginas web, imágenes, videos y otros archivos. Una URL consta de varios componentes, incluyendo el esquema (como "http://.º "https://"), el nombre de dominio (como "www.ejemplo.com"), la ruta del recurso y otros parámetros opcionales.

Sin embargo, a medida que la web ha crecido en tamaño y complejidad, el enfoque

[2]Cabe destacar que en esta época, los documentos eran principalmente texto plano, sin formato, y no existía la posibilidad de incluir imágenes o vídeos.

de direccionamiento basado en la ubicación física de los servidores puede presentar limitaciones. Por ejemplo, si un recurso se encuentra en una URL específica y esa URL cambia o el servidor deja de estar disponible, el acceso al recurso se verá comprometido.

IPFS aborda este problema mediante el uso de un sistema de direccionamiento basado en el contenido, en lugar de la ubicación. En IPFS, cada archivo y bloque de datos se identifican mediante su contenido, utilizando una función hash criptográfica. Esto permite que los archivos y bloques se puedan encontrar y acceder de forma fiable, independientemente de su ubicación física.

Esto permite a IPFS ofrecer una serie de ventajas sobre el sistema de direccionamiento basado en la ubicación de la web tradicional, como la resistencia a la censura, la persistencia de los datos y la verificabilidad del contenido. En la siguiente sección se profundizará en estas ventajas y en cómo IPFS las hace posibles.

2.2. IPFS como alternativa a HTTP

2.2.1. Introducción

IPFS fue presentado al mundo en 2014 por Juan Benet, en su informe técnico (whitepaper) *IPFS - Content Addressed, Versioned, P2P File System*[**?**]. Benet presenta el concepto de IPFS y su proposición de crear un sistema de archivos distribuido y descentralizado que permita a los usuarios almacenar y compartir archivos de forma segura y confiable.

Benet es también el fundador de Protocol Labs[**?**], una empresa dedicada a la creación de protocolos de código abierto para la Web3. IPFS es un proyecto de código abierto y pese a que Protocol Labs está detrás de este, no es el único contribuidor a su desarrollo. Esto es otro de los puntos fuertes de IPFS, la comunidad que lo rodea. En la sección **??**: '**??**' se profundiza en este aspecto.

En IPFS, cada archivo se identifica de manera única a través de su contenido mediante un hash criptográfico. Esto significa que cualquier nodo en la red puede actuar como un proveedor de contenido al almacenar y compartir archivos, permitiendo una mayor disponibilidad y un internet verdaderamente descentralizado. En lugar de depender de un único servidor web para acceder a un recurso, los usuarios pueden obtener el contenido de cualquier nodo que tenga ese recurso en particular.

Estos identificadores de contenido se conocen como CID (Content Identifier). Dado que un CID es un puntero que señala a un contenido particular, se puede usar un CID en vez de URL en un enlace. De esta manera se puede acceder a un recurso de manera fiable, independientemente de su ubicación física, mientras haya algún otro nodo de la red que el contenido que buscamos.

2.2.2. Fundamentos

IPFS opera a través de tres principios fundamentales que marcan una diferencia significativa con respecto a los sistemas de archivos convencionales: direccionamiento por contenido, red peer-to-peer y el grafo acíclico dirigido de Merkle (Merkle DAG).

Direccionamiento por Contenido: En IPFS, los archivos no se ubican por su dirección sino por su contenido. Cada archivo posee un identificador único, denominado CID (Content Identifier), generado a partir de un hash criptográfico de su contenido. Esta característica asegura la inmutabilidad de los archivos, es decir, los archivos no pueden ser alterados sin modificar su CID. Adicionalmente, el direccionamiento por contenido favorece la deduplicación, dado que archivos con contenido idéntico compartirán el mismo CID, lo que conlleva a su almacenamiento único dentro de la red.

Red Peer-to-Peer: IPFS se basa en una red descentralizada en la que cada integrante, o nodo, puede interactuar directamente con cualquier otro nodo, sin la necesidad de intermediarios o servidores centrales. Los nodos funcionan tanto como proveedores como consumidores de contenido, guardando y compartiendo fragmentos de archivos con otros nodos. Esta red peer-to-peer hace que el contenido sea más accesible y resistente a la censura, al evitar la existencia de un único punto de fallo o control.

Grafo Acíclico Dirigido de Merkle (Merkle DAG): Los archivos y sus relaciones dentro de IPFS se representan mediante una estructura de datos conocida como Merkle DAG. Un Merkle DAG es un grafo donde cada nodo tiene un identificador único (CID) que se genera a partir de su contenido y el de sus nodos hijos. Los nodos pueden ser hojas o nodos intermedios, dependiendo de si tienen o no nodos hijos. Los nodos hoja contienen datos binarios de los archivos, mientras que los nodos intermedios contienen enlaces a otros nodos. Los nodos intermedios permiten dividir archivos grandes en bloques más pequeños y formar estructuras jerárquicas, como directorios o sistemas de archivos. El Merkle DAG facilita la verificación de integridad y autenticidad de los archivos, dado que cualquier cambio en el contenido o en los enlaces se refleja en el CID del nodo afectado y sus ancestros.

Cada uno de estos conceptos se profundizará dentro del apartado correspondiente a continuación.

2.2.3. Arquitectura

IPFS es un conjunto de protocolos de código abierto que combina múltiples conceptos existentes de redes peer-to-peer (P2P), datos enlazados y otras áreas para permitir que los participantes intercambien fragmentos de archivos.

Estos conceptos concretados en protocolos forman distintos niveles de abstracción, cada uno de los cuales se puede utilizar de forma independiente y conforman la arquitectura de IPFS, también conocido como el *stack* de protocolos de IPFS. Esta pila de protocolos tiene cierto parecido al modelo OSI (Open Systems Interconnection), que es un modelo conceptual que caracteriza y estandariza las funciones de comunicación de un sistema de comunicación o red de computadoras, dividiéndolo en siete capas.

Cada capa se encarga de un aspecto específico de la comunicación. La figura **??** muestra el modelo OSI y las funciones de cada capa.

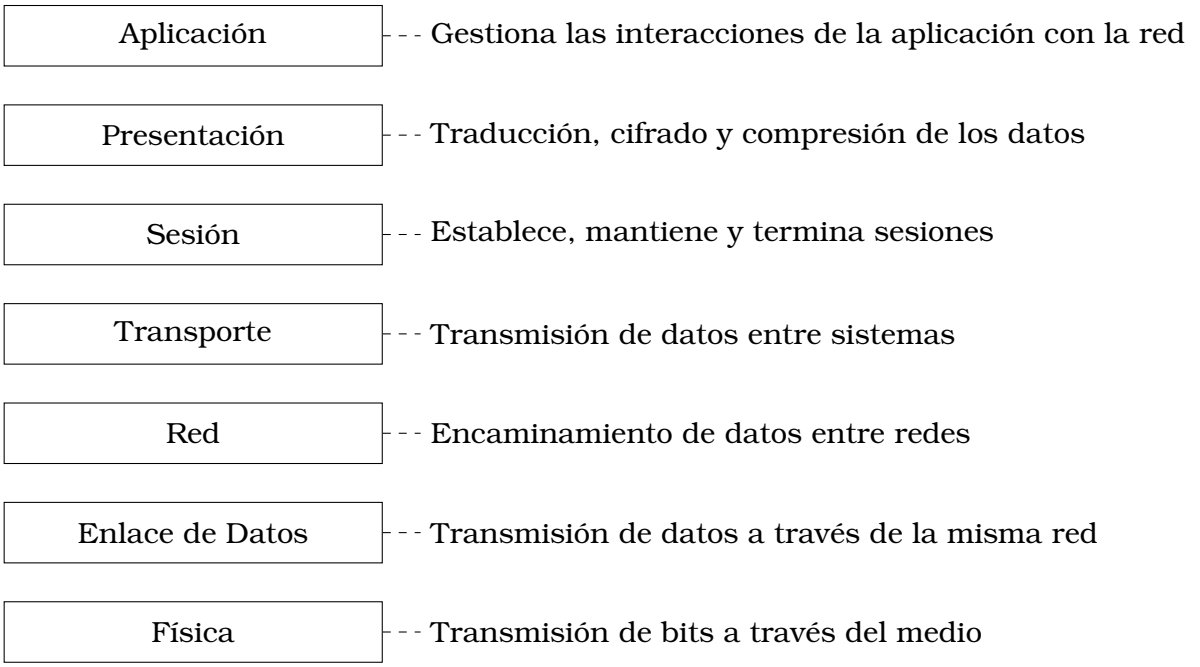

Figura 2.3: Modelo OSI

El stack de IPFS aunque parecido no es exactamente igual, pero el modelo OSI es una buena referencia para entender el de IPFS. En la figura **??** se muestra el stack de protocolos de IPFS.

Como se puede observar, este stack podría subdividirse en tres grupos según la funcionalidad que brinda cada capa.

2.2.4. Modelo de datos

2.2.5. Distribución de contenido

2.3. Ecosistema en torno a IPFS

2.3.1. Introducción

2.3.2. Proyectos basados en IPFS

2.3.3. Herramientas y librerías de IPFS

2.3.4. Comunidades en torno a IPFS

2.3.5. Integraciones de IPFS

Capítulo 3

Desarrollo de IPFShare

3.1. Casos de uso

3.2. Objetivos

3.3. Requisitos

3.4. Tecnologías

3.4.1. Tecnologías propuestas

3.4.2. Tecnologías usadas

3.5. Arquitectura del sistema

3.6. Implementación

3.6.1. Backend (Electron)

3.6.2. Frontend (React)

Capítulo 4

Resultados y conclusiones

Capítulo 5

Resultados y conclusiones

5.1. Resultados

hola

Capítulo 6

Trabajos futuros

Anexo

Característica	**IP/TCP**	**OSI**	**X.25**	**SNA**
Modelo	Suite de protocolos	Modelo de referencia	Protocolo de enlace	Suite de protocolos
Capas	4 (TCP/IP)	7	3	7
Año de lanzamiento	1974 (TCP) / 1981 (IP)	1984	1976	1974
Enfoque	Conmutación de paquetes	Conmutación de paquetes y circuitos	Conmutación de circuitos	Conmutación de paquetes y circuitos
Estándar	IETF	ISO	CCITT (ahora ITU-T)	IBM
Orientación	Red global	Interoperabilidad	Redes de área amplia (WAN)	Redes empresariales
Funcionalidades	Transmisión de datos, enrutamiento, control de flujo, control de congestión, conexión y desconexión	Transmisión de datos, enrutamiento, control de flujo, control de congestión, conexión y desconexión, servicios de presentación y aplicación	Transmisión de datos, control de flujo, conexión y desconexión	Transmisión de datos, enrutamiento, control de flujo, control de congestión, conexión y desconexión, servicios de presentación y aplicación
Uso en los años 90	Muy popular, base del Internet	Intento de reemplazar a TCP/IP, pero fracasó en la adopción generalizada	Utilizado en redes de área amplia (WAN), especialmente en Europa	Utilizado en redes empresariales, especialmente en sistemas mainframe de IBM

Descripción	Un modelo que se basa en la suite de protocolos TCP/IP para transmitir datos por Internet. El modelo es más simple y flexible que el modelo OSI y se usa ampliamente en la actualidad.	Un modelo que se basa en la suite de protocolos OSI para estandarizar la comunicación entre sistemas abiertos. El modelo segmenta múltiples funciones que el modelo IPTCP agrupa en capas únicas y define los servicios e interfaces para cada capa.	Un modelo que se basa en la suite de protocolos X.25 para proporcionar una conexión virtual entre terminales y computadoras a través de una red pública de conmutación de paquetes. El modelo fue uno de los primeros en ofrecer una comunicación confiable entre dispositivos remotos, pero ha sido reemplazado por tecnologías más rápidas y eficientes como Frame Relay e IP.	Un modelo que se basa en la suite de protocolos SNA para integrar los recursos informáticos distribuidos en una red jerárquica. El modelo fue desarrollado por IBM para conectar sus sistemas mainframe y periféricos, pero ha perdido popularidad frente a los modelos basados en IP.

Cuadro 1: Comparación de IP/TCP, OSI, X.25 y SNA en los años 90. Fuentes: [**?**], [**?**], [**?**], [**?**] [**?**] [**?**]