

TEMA 109. LA RED INTERNET Y LOS SERVICIOS BÁSICOS

Actualizado a 20/04/2023

1. HISTORIA DE INTERNET

- 1969 → Aparece ArpaNET (Advanced research projects agency NETwork) desarrollada en el MIT para ARPA. Usa protocolo NCP.
- 1972 → Evolucion a DARPA. Nace el InterNetworking Working Group.
- 1983 → ARPANET adopta TCP/IP. Empieza a llamarse Internet y a tener estándares propios.
- 1984 → Empieza a crearse NSFNet (National Science Foundation Network). Interconecta centros de supercómputo e instituciones académicas. Adopta TCP/IP. Cambia arquitectura -> De redes agregadas a jerárquicas basada en 1 backbone.
- 1988: Nace en España la Red IRIS (fines académicos y de investigación), gestionada por Red.es.
- 1989 → Desaparece ArpaNet, que fue sustituida por la existente NSFNet.
- 1992 → Se crean ISOC y RIPE.
- 1995 → NFSNet deja de ser el backbone de Internet. Sustituido por puntos de interconexión comerciales a la red Internet → NAP.
- 1995 → Uso comercial de internet (ISP's, Internet Service Providers).

2. ORGANIZACIONES DE INTERNET

2.1. ICANN

La Internet Corporation for Assigned Names and Numbers (ICANN) es una organización sin fines de lucro responsable de asignar las **direcciones IP**, así como la gestión y administración de los **nombres de dominio de primer nivel genéricos (gTLD)** y de **códigos de países (ccTLD)**. También se encarga de asignar los **puertos** en Internet, que son precisos a la hora de especificar los estándares.

2.1.1. IANA

La Internet Assigned Numbers Authority (IANA) es la entidad que supervisa la **asignación global de direcciones IP, sistemas autónomos, servidores raíz de nombres de dominio DNS y otros recursos** relativos a los protocolos de Internet. IANA coordina a nivel global el "**DNS root**". IANA gestiona el dominio .int (top-level). Asimismo, IANA administra el dominio .ARPA en estrecha relación con el IAB (Internet Architecture Board). **IANA NO asigna las direcciones IP, pero si la supervisa.** IANA es también la entidad **encargada de la coordinación de los Números de Sistema Autónomo** (Autonomous System Numbers - ASN) usados por varios protocolos de enrutamiento. IANA actualmente es un departamento operado por ICANN.

2.2. W3C

Elabora estándares abiertos para la interoperabilidad en la Web. Ha elaborado los estándares de HTML, XML, XHTML, SOAP, WSDL, CSS, WAI, DOM...

2.2.1. SERVICIOS WEB

WSDL: Web Services Description Language. Está basado en XML. Se utiliza para describir servicios web (WS). Es un estándar del W3C.

SOAP: Simple Object Access Protocol. También es el acrónimo de Services Oriented Architecture Protocol. SOAP es un protocolo para hacer llamadas a métodos de objetos remotos. Es un estándar del W3C.

UDDI: Universal Description, Discovery and Integration. Centraliza servicios en un registro común tal que los servicios web de la red puedan publicar su localización y descripción, y hace que sea fácil descubrir que servicios están disponibles en la red. Es un estándar de OASIS.

- Páginas blancas: Dirección, contacto y otros identificadores conocidos.
- Páginas amarillas: Categorización industrial basada en taxonomías.
- Páginas verdes: Información de servicio. Información técnica sobre los servicios que aportan las propias empresas.

REST: Representational state transfer. Es una arquitectura ligera de llamada a servicios remotos, basada típicamente en HTTP, alternativa a Web Services.

2.3. ISOC

Internet Society (ISOC) es una organización no gubernamental y sin ánimo de lucro, constituida como la única organización dedicada exclusivamente al desarrollo mundial de Internet. Facilita la **cooperación y coordinación** en los asuntos relacionados de Internet. De esta organización dependen el IETF e IAB. Dentro de IETF, el IESG (Grupo Directivo de Ingeniería de Internet, IESG por su sigla en inglés) se encarga de Coordinación entre organizaciones. IAB supervisa y aprueba las normas. Ejerce sus funciones a través de:

- **IAB** (Internet Architecture Board) → **Supervisión y aprobación de normas.**
- **IETF** (Internet Engineering Task Force) → **Especificación de estándares.**
- **IESG** (Internet Engineering Steering Group) → **Coordinación.**

2.3.1. IAB

Acrónimo del organismo responsable de definir toda la arquitectura de Internet [GSI2010]. Depende de ISOC.

IRTF: Internet Research Task Force (IRTF). Para la investigación de protocolos, arquitecturas, tecnologías y servicios. Depende del IAB.

2.3.2. IETF

IETF (Internet Engineering Task Force) se encarga de Especificación de estándares (RFCs). Depende de ISOC.

2.3.3. IESG

El Internet Engineering Steering Group (o Grupo Directivo de Ingeniería de Internet, IESG por su sigla en inglés) es un organismo compuesto por el presidente del Internet Engineering Task Force (IETF) y directores de área. Proporciona la revisión técnica final de los estándares de Internet y es responsable de la gestión del día a día de la IETF. Recibe las apelaciones de las decisiones de los grupos de trabajo, y el IESG toma la decisión de avanzar documentos en el desarrollo de las normas.

2.3.4. RFCS (REQUEST FOR COMMENTS)

Elaborados por el IETF (Internet Engineering Task Force). Cualquier persona puede enviar propuestas para ser RFC (Request For Comments). Cada RFC tiene un título y un número asignado, que no puede repetirse ni eliminarse aunque el documento se quede obsoleto. Los RFCs se han publicado tradicionalmente en texto ASCII. El estado de una RFC creada por el IETF tiene que ver con las diferentes fases a seguir en su definición: protocolo estándar, borrador estándar, propuesta estándar, experimental, informativo y protocolo histórico.

2.4. RIRS

Son los asignadores regionales de Internet. Asignan las direcciones IP.

- AfriNIC: Africa.
- APNIC: Asia y el Pacífico
- ARIN: America del Norte
- LACNIC: Latinoamérica y Caribe
- RIPE NCC: Europa, Oriente Medio y Asia Central.

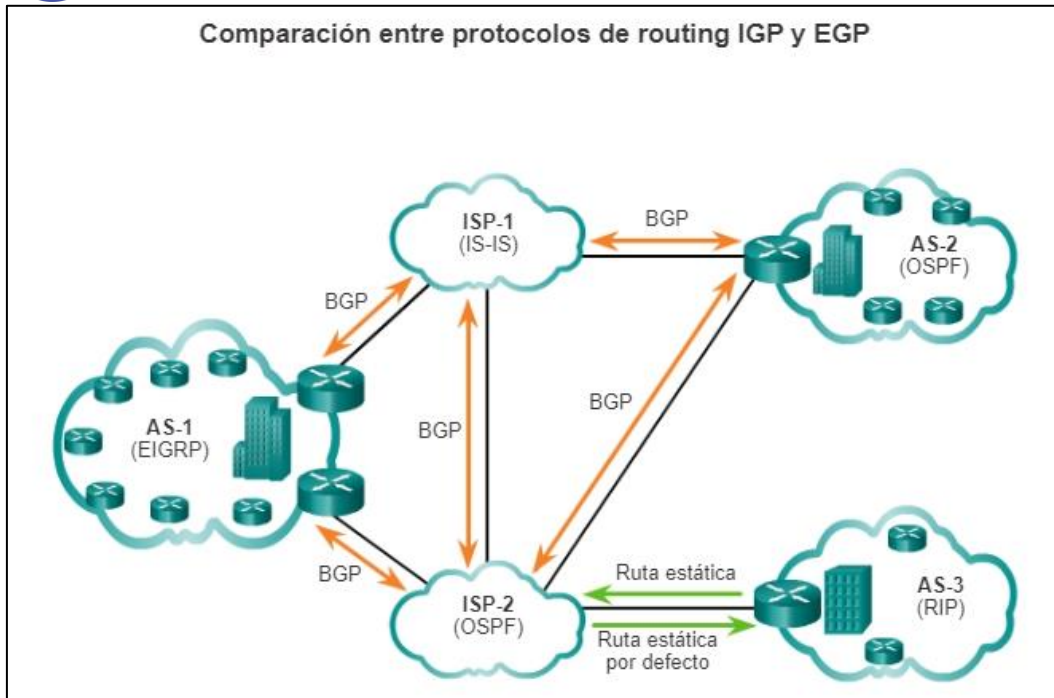
En España se gestionan los dominios .es por la entidad pública empresarial **Red.es**. El punto neutro español es la organización sin ánimo de lucro **Espanix**.

3. CONCEPTOS BÁSICOS

3.1. ENRUTAMIENTO E CONEXIÓN EN INTERNET

A continuación se hace una somera descripción de algunos conceptos elementales que son necesarios para comprender la arquitectura de alto nivel de Internet.

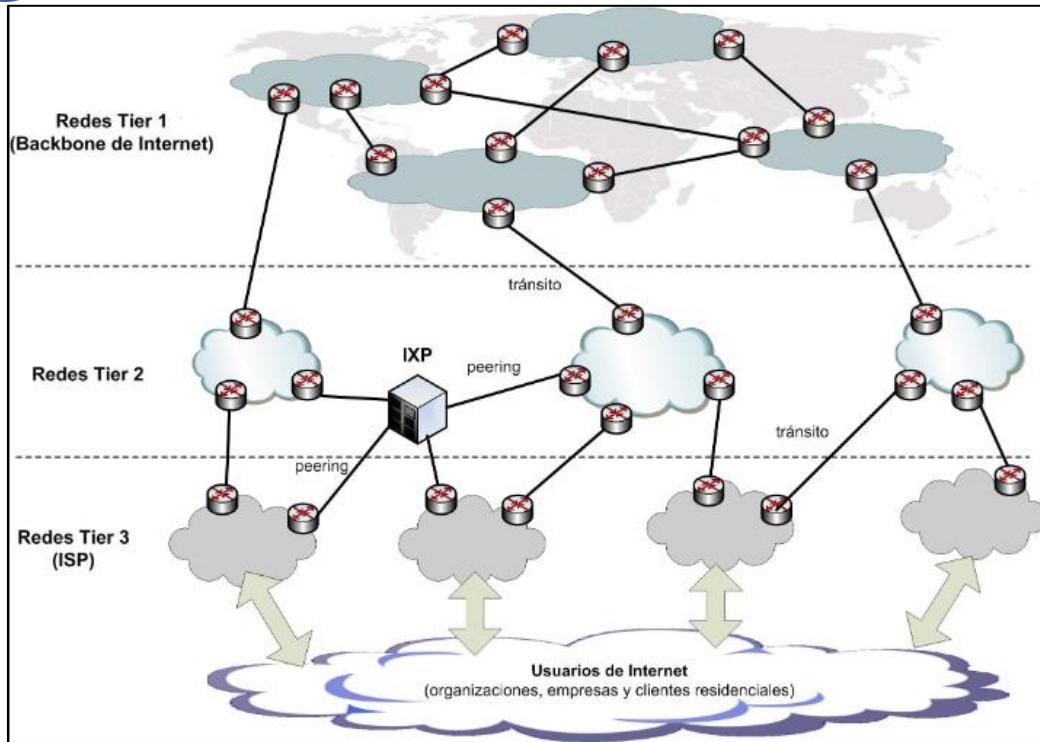
- **Autonomous System (AS):** Es una red o un grupo de redes IP bajo el control de un único ISP. Los AS son asignados por la IANA en bloques a las autoridades regionales RIR, y son estas últimas las que los asignan a los ISPs. A cada AS se le asigna un número de sistema autónomo (ASN).
- **Interior Gateway Protocol (IGP):** Se utiliza para el routing dentro de un AS. También se lo denomina "routing interno de AS". Las empresas, las organizaciones e incluso los proveedores de servicios utilizan un IGP en sus redes internas. Los IGP incluyen RIP, IGRP, EIGRP, OSPF e IS-IS.
- **Exterior Gateway Protocol (EGP):** Se utiliza para el enrutamiento entre AS. Los proveedores de servicios y las empresas grandes pueden interconectarse mediante un EGP. El protocolo de gateway fronterizo (**BGP**) es un ejemplo de EGP y es el protocolo de routing utilizado por Internet.
- **Border Gateway Protocol (BGP):** Es el EGP más popular. BGP es un protocolo mediante el cual los ISPs intercambian información de encaminamiento y rutas entre sistemas autónomos. Este intercambio de información de enrutamiento se hace entre los routers externos de cada sistema autónomo, los cuales deben soportar BGP y se utiliza el protocolo BGP para el intercambio de información entre los AS.



3.2. JERARQUÍA EN INTERNET: TIERS

La estructura de Internet está basada en la interconexión de redes de forma jerárquica mediante una clasificación de niveles, llamados Tiers.

- Tier 1: Red troncal de internet (Backbone), Grandes operadores globales (global carriers).
- Tier 2: Redes de proveedores de tránsito y acceso internacional.
- Tier 3: Redes de proveedores de acceso local (ISPs).
- Tier 4: Redes corporativas.



Peering: Enlace directo entre AS (Autonomous Systems) por el que transmiten tráfico sin coste:

- Tier 1. Solo se accede por acuerdos de peering.
- Tier 2. Puede tener acuerdos de peering. Deben pagar el tráfico para alcanzar Tier 1 y poder comunicarse con todo el mundo.
- Tier 3. Solo se conectan a Tier 2. Siempre pagan tráfico en Internet.

3.3. PROTOCOLOS

IPv4 (Internet Protocol v4) → Define direcciones IP de 32 bits. Es la versión inicialmente usada en Internet, y es todavía el protocolo dominante. El gran crecimiento de Internet provocó el casi agotamiento de las direcciones IPv4.

IPv6 (Internet Protocol v6) → Define direcciones IP de 128 bits. Nacida como consecuencia del agotamiento de direcciones IPv4, incorpora también mejoras sobre el protocolo para mejorar el rendimiento, seguridad y gestión. En principio IPv6 no es compatible con IPv4, si bien pueden existir infraestructuras de interoperabilidad entre ambas.

DNS (Domain Name System) → Sistema de nombres de dominio, encargada de la traducción de nombres de dominio a direcciones IP y viceversa.

3.4. JERARQUÍA EN INTERNET

El enrutamiento en Internet se realiza por medio del protocolo **BGP (Border Gateway Protocol)**, que permite intercambiar información de rutas y números de AS entre los ISPs para poder establecer las rutas óptimas.

Conceptos de enrutamiento:

- **AS (Autonomous System):** Designa un grupo de redes IP bajo el control de un ISP. Los AS son asignados por la IANA en bloques a las autoridades regionales RIR, y son estas últimas las que los asignan a los ISPs.
- **PoP (Point of presence):** Ubicaciones donde un ISP ofrece sus servicios de conexión a Internet a los clientes. El cliente, mediante alguna de las tecnologías de acceso (siendo la más extendida ADSL), se conecta al PoP local del ISP y a través del mismo a Internet.
- **NAP (Network Access Point)/ IXP (Internet Exchange Point):** Punto de interconexión entre ISP, a través del cual los ISP intercambian tráfico entre sus sistemas autónomos por medio de acuerdos de peering. Estos puntos evitan tener que enviar tráfico a una entidad de un *tier* superior, que tiene un coste superior.
- **Peering:** Intercambio de tráfico IP entre dos ISP.
- **ISP (Internet Service Provider):** Organización que conecta usuarios finales o negocios a Internet.

3.4. PROTOCOLO DNS

El protocolo DNS (*Domain Name System*) permite resolver los nombres de dominios y de hosts en sus direcciones IP. Consiste en una **base de datos global, distribuida y jerárquica** que contiene registros agrupados por zonas. Funciona sobre los puertos **53 udp o tcp**.

Los nombres de dominio tienen una estructura jerárquica, siendo el primer nivel establecido por la ICANN. En cada país hay organismos que regulan y establecen las condiciones para la concesión de un dominio. Estos organismos generalmente designan en agentes el trámite de la concesión del dominio. En España **Red.es** tiene encomendada la autoridad de registro de los nombres de dominio de Internet bajo el indicativo de primer nivel correspondiente al país de España (.es).

El protocolo es cliente/servidor. Los servidores DNS son:

- Root name server (nivel más alto).
- Authoritative server (responsable de una zona).
- Caching name server (acelera consultas repetidas).

Cada dominio tiene por lo menos un **servidor DNS autoritativo**, que publica información sobre el dominio y los servidores de nombre de cualquier dominio subordinado. Los elementos superiores de la jerarquía son los servidores de nombre raíz, a quien se consultará en el caso de resolver un nombre de dominio de primer nivel (*TLD – Top Level Domain*).

SECOND LEVEL DOMAIN (SLD): segundo nivel o nivel empresarial de nombres de dominio. Ejemplo: si tenemos <http://red.empresa.com/>, el TLD=.com y el SLD= empresa.com (se incluye el TLD) Un servidor DNS autoritativo únicamente responde a consultas DNS de datos de los que se ha configurado como fuente original. En contraposición, pueden existir servidores DNS no autoritativos, que reenvían todas las consultas que reciben a otro servidor, manteniendo simplemente una caché de consultas.

La forma de operación habitual para resolver nombres DNS suele ser:

1. Consultar fichero “hosts”
2. Consultar el servidor más prioritario (de forma manual o por DHCP). Si no está disponible acude por orden a la lista de servidores DNS configurados.
3. El servidor DNS responde si tiene la información cacheada o es “Authoritative” de la zona correspondiente. Si no, puede redirigir la consulta al de orden superior.
4. Para hacer llegar la respuesta al cliente se sigue el camino inverso a las consultas.

Tipos de registro DNS más comunes:

- **SOA** (Start of Authority): Información sobre la zona DNS, especialmente en cuanto a transferencias de zona.
- **A** (Address): Contiene una IPv4.
- **AAAA** (Address): Contiene una IPv6
- **MX** (Mail eXchanger): Referencia un servidor de correo para el reenvío de mensajes.
- **NS** (Name Server): Servidores de nombre.
- **CNAME**: Alias de nombres de dominio.
- **CAA** (Certification Authority Authorization): Permite limitar las autoridades certificadoras válidas para un dominio.
- **PTR** (Pointer): Registro inverso, funciona de manera opuesta al A, traduce IPs a nombres de dominio.
- **TXT** (Text): registro de texto. Permite insertar el texto que desees. Suele utilizarse para verificar la autoridad de dominio o evitar usos no correctos de las direcciones de correo.

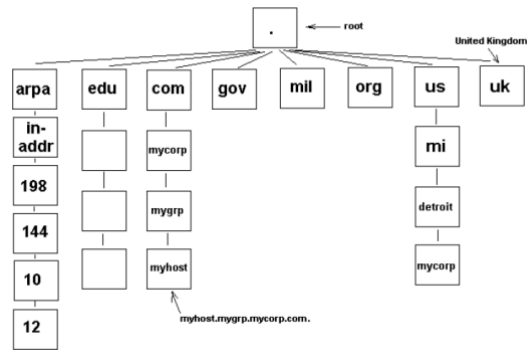
Existe también el Multicast DNS, que no usa servidores y se utiliza para redes pequeñas.

Los nombres de dominio

Tienen estructura jerárquica, siendo el primer nivel (Top Level Domain) establecido por ICANN. En cada país hay organismos que gestionan las concesiones, siendo en España Red.es.

Los niveles son:

- TLD (Top Level Domain). Parte final del dominio. Tipos:
 - GTLD (General TLD). Genéricos, con tres o más letras de largo. Los más comunes:
 - edu (instituciones educativas)
 - gov (instituciones gubernamentales)
 - com (organizaciones comerciales)
 - mil (grupos militares)
 - net (redes)
 - int (organizaciones internacionales)
 - org (otras organizaciones)
 - Se han ido creando otros como .aero (espacio), .biz (negocios), .coop (cooperativas), .info (uso genérico), .museum, .name (genérico), .pro (profesionales), .jobs (trabajo), .int, .mobi, .mil (ejército de EEUU), .tel (servicios de telefonía), .travel (viajes)...
 - CCTLD (Country Code TLD): Usados por un país o un territorio dependiente. Dos letras de longitud. Más de 200. Ejemplos son .es, .uk, .fr
- Resto de niveles con significado dependiendo de la situación. Ejemplos:
 - SLD (SECOND LEVEL DOMAIN): segundo nivel o nivel empresarial de nombres de dominio
 - Third Level: tercer nivel. Hace referencia a un servicio dentro de la empresa



Servicio whois

Protocolo abierto TCP (puerto 43), basado en petición respuesta para efectuar consultas en una base de datos obteniendo el propietario de un nombre de dominio o dirección IP.

3.5. SEGURIDAD

Normas técnicas elaboradas por la Internet Engineering Task Force (IETF) para mejorar la seguridad de la infraestructura de Internet:

- IPsec (Internet Protocol Security). Seguridad extremo a extremo en la capa IP
- TLS (Transport Layer Security). Seguridad de las comunicaciones sobre Internet
- Kerberos Network Authentication System. Verificación de identidad en una red abierta.
- DNSSEC (Domain Name System Security Extensions). Integridad y autenticidad respuestas DNS.
- DANE (DNS-based Authentication of Named Entities). Complementa DNSSEC- conexión criptográficamente segura.

Normas técnicas W3C para mejorar la seguridad:

- Content Security Policy.
- Cross-Origin Resource Sharing.
- XML Signature, XML Encryption and related specifications.
- Cryptographic APIs for JavaScript.

Estándares de seguridad de OASIS:

- DSS. Digital Signature Services. Firma digital para XML.
- KMIP. Key Management Interoperability Protocol. Tecnologías de clave con cifrado asimétrico.
- SAML. Security Assertion Markup Language. Creación e intercambio de información de seguridad.
- XACML. eXtensible Access Control Markup Language. Políticas de control de acceso.

Elementos de red: Firewalls, IDS, Servidores de Acceso, etc.

4. SERVICIOS DE INTERNET

4.1. WWW (WORLD WIDE WEB)

Es una colección global de documentos, imágenes, multimedia, aplicaciones, y otros recursos interconectados a nivel lógico mediante hipervínculos y referenciados mediante URIs (*Uniform Resource Identifiers*), que proporcionan un sistema global de referencias de nombres.

HTTP (Protocolo de Transferencia de Hipertexto) es el principal protocolo de acceso a la www. Es un protocolo de nivel de aplicación para sistemas hipermedia distribuidos y colaborativos. Es la base de la comunicación en la www, en que los documentos de hipertexto incluyen hipervínculos a otros recursos a los que el usuario puede acceder fácilmente.

Fue desarrollada inicialmente por Tim Berners-Lee en 1989, si bien más tarde pasó a ser coordinada por el IETF.

Versiones de HTTP:

- **HTTP/1.1:** Introducida en el RFC 2068, más tarde reemplazado con el RFC 2616 y finalmente por el RFC 9112.
- **HTTP/2:** Revisión del protocolo para introducir un mejor rendimiento y menor latencia, mediante el uso de multiplexado de solicitudes por conexión tcp y la compresión de cabeceras entre otros mecanismos. Puede utilizarse sobre TLS 1.2 mediante la extensión ALPN. Dispone de un mecanismo de negociación de la versión HTTP para mantener la compatibilidad con el anterior. Su soporte por navegadores y servidores web es ya amplio. La RFC más reciente que define el protocolo HTTP/2 es la RFC 9113.
- **HTTP/3:** Nueva revisión buscando mejoras de rendimiento. En lugar de utilizar TCP para la capa de transporte, funciona sobre QUIC, un protocolo de capa de transporte desarrollado inicialmente por Google, que a su vez funciona sobre **UDP**. El soporte por los navegadores de esta versión es todavía limitado. La RFC más reciente que define el protocolo HTTP/3 es la RFC 9114.

HTTP proporciona múltiples tipos de mecanismos de autenticación, los más comunes son:

- Autenticación básica: El usuario y clave se envían codificadas en base64 al servidor.
- Autenticación “digest”: incluye un mecanismo desafío-respuesta, las credenciales nunca se envían en claro.

Métodos HTTP:

- **GET:** solicita una representación del recurso especificado.
- **HEAD:** Igual que GET, pero la respuesta no incluye el cuerpo del mensaje. Es útil para consultar metadatos del recurso especificado.
- **POST:** Solicita que el servidor acepte la nueva entidad contenida en la petición al recurso especificado en la URI.
- **PUT:** Solicita que la entidad contenida en la petición sea almacenada en la URI especificada. Si la URI se refiere a un recurso existente, debe ser modificado; si no existe debe crearse.
- **DELETE:** Elimina el recurso especificado.
- **TRACE:** Devuelve la solicitud recibida, de forma que un cliente pueda comprobar si se han realizado cambios o adiciones por los servidores intermedios.
- **OPTIONS:** Devuelve los métodos HTTP que el servidor soporta para la URL especificada.

- **CONNECT:** Convierte la solicitud de conexión en un túnel TCP/IP, generalmente para establecer un túnel SSL cifrado a través de un proxy HTTP.
- **PATCH:** Aplica modificaciones parciales a un recurso.

Códigos de resultado:

- 1xx: Mensajes informativos.
- 2xx: Correcto (Successful)
- 3xx: Redirección
- 4xx: Error de cliente
- 5xx: Error del servidor

Cookies → Las cookies HTTP son pequeños conjuntos de información almacenados en el ordenador del usuario por el navegador, mientras se navega por un sitio web. Se diseñaron como un mecanismo para mantener información de estado (como los artículos en el carrito de la compra de una tienda online) o la actividad de navegación del usuario (inicio de sesión, páginas visitadas, etc.). También puede usarse para mantener información de campos de formularios que el usuario rellenó anteriormente. Actualmente se emplea el estándar **RFC 6265** (están obsoletos los RFC 2109 y 2965 en que se definió anteriormente).

4.2. FTP

Intercambio de ficheros. Puede ser anónimo, o requerir identificación. Utiliza dos puertos: un puerto de control (**21 tcp**), y un puerto de datos (**20 tcp**) que puede realizarse de dos modos:

- **Modo Activo (En Servidor) (PORT):** El Servidor le ordena al Cliente que abra el puerto, y el Servidor se conecta a ese puerto (puede haber problemas con firewalls)
- **Modo Pasivo (En Servidor) (PASV):** El Servidor le indica al Cliente a que puerto debe conectarse. El Servidor abre el puerto en su equipo, y espera que el Cliente se conecte al mismo. El modo pasivo es muy recomendable cuando el cliente tiene una conexión con cortafuegos, ya que en esta modalidad, tanto el canal de control como el canal de datos los abre el Servidor. El cliente se conectará a un puerto por encima de 1023 elegido al azar por el Servidor

Seguridad:

- **FTPs:** FTP sobre SSL. Por defecto puerto 990 para control, 989 datos.
- **sFTP:** transferencia sobre SSH. Puerto 22 por defecto.

4.3. MAIL

Intercambio de mensajería fuera de línea. Los servidores de mail aceptan, encaminan, entregan y almacenan mensajes. No se requiere que ni los usuarios ni otros servidores estén online de forma simultánea; deben conectarse para enviar o recibir los mensajes.

Los protocolos asociados más importantes son:

- **POP3 (Post Office Protocol):** Para obtener un mensaje de un servidor mediante su descarga. Puerto **110 tcp**, o **995 tcp** si es POP3 sobre SSL.
- **IMAP (Internet Message Access Protocol):** Para obtener los mensajes de un servidor, pero a diferencia de POP3 permite almacenar los mensajes en el servidor de correo. Puerto **143 tcp**, o **993 tcp** si es IMAP sobre SSL.

- **SMTP** (*Simple Mail Transfer Protocol*): Envío de los correos desde un cliente, o intercambio entre servidores para llegar al destino. Puerto **tcp 25**.
- **MIME** (*Multipurpose Internet Mail Extensions*): Extiende el formato de los mensajes de texto para soportar otros juegos de caracteres distintos al ASCII, así como adjuntos de audio, video, imágenes,...

4.4. OTROS SERVICIOS

- **Archie**: Permite localizar ficheros mediante índices y obtenerlos por FTP.
- **Finger**: Proporciona información sobre los usuarios de un servidor. Puerto 79 tcp.
- **WAIS**: Permite encontrar documentos indexados en BD mediante una clave y nombre de la BD.
- **Telnet**: Acceso remoto a una cuenta de usuario. Claves viajan en claro por la red. Puerto **23 tcp**.
- **Chat**: Comunicación escrita a través de Internet entre dos o más personas. Se escribe bajo pseudónimos o alias llamados Nick.
- **P2P**: Intercambio de ficheros entre pares. Nodos que se comportan a la vez como clientes y servidores.
- **SSH**: Acceso remoto (Protocolo Seguro), ha sustituido al protocolo telnet. Puerto **22 tcp**.
- **News**: Herramientas creadas para mantener discusiones entre los usuarios, sobre áreas temáticas definidas. Protocolos asociados: **NNTP, UUCP**.
- **TOR** (The Onion Router): Red distribuida y de baja latencia superpuesta sobre Internet (darknet o Deep web). Se consigue el anonimato de los usuarios (IP).
- **NTP** (Network Time Protocol): Sincronización de relojes.
- **Rlogin**. Acceso remoto a sistemas informáticos. Puerto tcp 513.
- **NFS**. Network File System (Sistema de archivos de red).
- **IIP**. Protocolo de envío y control de trabajos de impresión desde un navegador. Puerto tcp 631.
- **Buscadores, portales, comercio electrónico, sistemas de pago online** (como PayPal o Servicio MasterCard Secure Code y Verified by Visa), **weblogs, wikipedia, analítica web**.