



INSTITUTO TECNOLÓGICO SUPERIOR DE POZA RICA

INGENIERÍA SISTEMAS COMPUTACIONALES

Materia

Desarrollo en Internet

Nombre

Javier Alexis Mendoza Gracia

Docente

Simón García Ortiz

POZA RICA DE HGO, VER.

25 de septiembre del 2020

Introducción

Como introducción acerca de este tema tenemos que, la tecnología es un conjunto de soluciones que nos permite observar crear Web está basada en el protocolo HTTP para la transmisión de hipertexto sobre protocolos de comunicaciones TCP/IP, sirve para poder proporcionar una nueva visión de internet.

Si antes, con servicios como FTP, la Red se concebía como un conjunto de nodos servidores cuya topología y estructura física había que conocer, ahora se logra un nivel de abstracción mayor: lo que importa es la información, los contenidos en términos del WWW, en lugar de su ubicación, de las características de los servidores en los que se encuentra ubicada o de los dispositivos empleados para acceder a la misma: computadora personal, PDA , WebTV, teléfono móvil, etc.

Introducción a las tecnologías web.

Perspectiva histórica del Internet

El objetivo era crear una red de comunicaciones militar capaz de seguir funcionando después de un ataque nuclear.

Los primeros pasos fueron:

En 1964 Paúl Baran, de Rand Corporation propuso la utilización de una red de conmutación de paquetes.

En 1968 se puso en funcionamiento el NPL (National Physics Laboratory) en el Reino Unido.

En 1969 se creó ARPANET, una red para comunicar los distintos centros implicados en el proyecto ARPA.

Los miembros de ARPANET fueron:

Stanford Research Institute (SRI).

Universidad de California, Santa Barbara.

Universidad de California, Los Ángeles (UCLA).

Universidad de Utah.

Evolución de ARPANET:

1973. Primera conexión internacional con Inglaterra y Noruega.

Surge el protocolo TCP/IP para estandarizar la comunicación entre redes.

1986. Surge NSFNET (National Science Foundation Network) para interconectar 5 grandes centros de supercomputación.

Miembros de NSFNET:

Princeton

Pittsburgh

UCSD

UIUC

Cornell

Evolución:

1990. Desmantelamiento de ARPANET.

1992. Tim Berners-Lee, físico del CERN European Organization for Nuclear Research (Suiza), inventa la Word Wide Web (WWW).

1993. Surge InterNIC.

1995. Sun Microsystems presenta un nuevo lenguaje de programación (Java).

Internet nació a partir de una red denominada ARPANET, diseñada y desarrollada en 1969 para el Departamento de Defensa de Estados Unidos, creada para mantener la comunicación entre computadoras en caso de guerra. Según el Pentágono suponía un gran peligro el hecho de que

todas las comunicaciones dependieran de un gran centro neurálgico que podría ser destruido con una sola bomba, quedando así sin posibilidades ante el enemigo. Debido a todas estas circunstancias, la red fue diseñada sin centro neurálgico.

Estados Unidos fue capaz de desarrollar una red que funcionara (la antecesora de la actual Internet) y los usuarios académicos e investigadores que tenían acceso a ella rápidamente se volvieron adictos. La demanda por la red muy pronto se esparció. Los desarrolladores de Internet en Estados Unidos, el Reino Unido y Escandinavia, en respuesta a las presiones del mercado, empezaron a poner el software de IP (Internet Protocol) en todo tipo de computadoras. Se llegó a convertir en el único método práctico para comunicar computadoras de diferentes fabricantes. Al mismo tiempo que Internet se consolidaba, al mismo tiempo, muchas compañías y otras organizaciones empezaron a construir redes privadas usando los mismos protocolos de ARPAnet. Parecía obvio que si estas redes podían comunicarse entre sí, los usuarios de una red podrían comunicarse con usuarios de otra y todo el mundo sería beneficiado.

De estas nuevas redes, una de las más importantes fue la NSFNET, auspiciada por la Fundación Nacional de la Ciencia (National Science Foundation), una agencia del gobierno de Estados Unidos. Al final de los ochenta la NSF creó cinco centros de supercómputo en universidades

importantes. Hasta ese entonces, las computadoras más rápidas del mundo sólo estaban a disposición de los fabricantes de armamento y de algunos investigadores de compañías muy grandes. Con la creación de centros de supercómputo, la NSF ponía éstas a disposición de cualquier investigación escolar. Al principio, la NSF trató de utilizar la red ARPAnet para la comunicación de los centros, pero esta estrategia falló debido a problemas burocráticos. En respuesta a esto, la NSF decidió construir su propia red basada en la tecnología IP de ARPAnet, ésta red conectaba los centros mediante enlaces telefónicos de 56,000 bits por segundo. El costo de la línea telefónica depende de la distancia por esta razón, se decidió crear redes regionales. En cada región del país las escuelas podían conectarse a su vecino más cercano. Cada cadena estaba conectada a un centro de supercómputo en un solo punto.

El hecho de compartir supercomputadoras permitió a los centros de cómputo compartir recursos no relacionados con los centros. Repentinamente, las escuelas que participaban en la red contaron

con un amplio universo de información y colaboradores al alcance de sus manos. El tráfico en la red se incrementó con el tiempo hasta que las computadoras que la controlaban y las líneas de

teléfono conectadas a ellas se saturaron. . En 1987 se celebró un contrato para administrar y

actualizar la red, con la compañía Merit Network Inc., en colaboración con IBM y MCI. La vieja red fue mejorada con líneas telefónicas de mayor velocidad (por un factor de 20) y con computadoras más poderosas.

Protocolo http

El Protocolo de Transferencia de HiperTexto (Hypertext Transfer Protocol) es un sencillo protocolo cliente-servidor que articula los intercambios de información entre los clientes Web y los servidores HTTP. Fue propuesto por Tim Berners-Lee, atendiendo a las necesidades de un sistema global dedistribución de información como el World Wide Web.

Desde el punto de vista de las comunicaciones, está soportado sobre los servicios de conexión TCP/IP, y funciona de la misma forma que el resto de los servicios comunes de los entornos UNIX: un proceso servidor escucha en un puerto de comunicaciones TCP (por default, el 80), y espera las solicitudes de conexión de los clientes Web. Una vez que se establece la conexión, el protocolo TCP se encarga de mantener la comunicación y garantizar un intercambio de datos libre de errores.

HTTP se basa en sencillas operaciones de solicitud/respuesta. Un cliente establece una conexión con un servidor y envía un mensaje con los datos de la solicitud. El servidor responde con un mensaje similar, que contiene el estado de la operación y su posible resultado. Todas las operaciones pueden adjuntar un objeto o recurso sobre el que actúan; cada objeto Web (documento HTML, Archivo multimedia o aplicación CGI) es conocido por su URL.

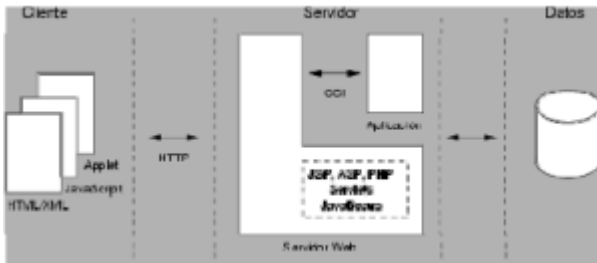
HTTP se utiliza para transmitir recursos, que incluyen además de archivos, el resultado de la ejecución de un programa, una consulta a una base de datos, la traducción automática de un documento, etc. Para un servidor HTTP, los recursos son o bien archivos, o bien el resultado de la ejecución de un programa.

Los tipos MIME (Multipurpose Internet Mail Extensions) son un estándar para el envío de información binaria a través de caracteres alfanuméricos. Este estándar permite que, a través del protocolo HTTP (que maneja información en modo texto), podamos transferir archivos no-textuales, como pueden ser imágenes, audio, vídeo, programas ejecutables etc. Los tipos MIME definen grupos (antes del carácter "/") y tipos (después del carácter "/"). Así el tipo MIME "text/html" define a todos los archivos de texto que contienen código HTML, el tipo "video/mpeg" define a todos los archivos de vídeo almacenados en formato mpeg, etc. Para indicar cualquier tipo se puede utilizar el carácter "*", tanto en el tipo como en el grupo. De este modo, el tipo MIME "image/*" representa a todos los archivos de imagen, ya estén almacenadas en formato gif, jpeg, bmp, etc.

Para profundizar más en el funcionamiento de HTTP, veremos primero un caso particular de una transacción HTTP; en los siguientes apartados se analizarán las diferentes partes de este proceso. Cada vez que un cliente realiza una petición a un servidor, se ejecutan los siguientes pasos:

- Un usuario accede a una URL, seleccionando un enlace de un documento HTML o introduciéndola directamente en el campo Locación del cliente Web.
- El cliente Web decodifica la URL, separando sus diferentes partes. Así identifica el protocolo de acceso, la dirección DNS o IP del servidor, el posible puerto opcional (el valor por default es 80) y el objeto requerido del servidor.
- Se abre una conexión TCP/IP con el servidor, llamando al puerto TCP correspondiente. Se realiza la petición. Para ello, se envía el comando necesario (GET, POST, HEAD,...), la dirección del objeto requerido (el contenido de la URL que sigue a la dirección del servidor), la versión del protocolo HTTP empleada y un conjunto variable de información, que incluye datos sobre las capacidades del browser, datos opcionales para el servidor,...

- El servidor devuelve la respuesta al cliente. Consiste en un código de estado y el tipo de dato MIME de la información de retorno, seguido de la propia información.
- Se cierra la conexión TCP.



Este proceso se repite en cada acceso al servidor HTTP. Por ejemplo, si se recoge un documento HTML en cuyo interior están insertadas cuatro imágenes, el proceso anterior se repite cinco veces, una para el documento HTML y cuatro para las imágenes.

Arquitectura del WWW

La arquitectura del WWW se puede ver desde dos puntos de vista Arquitectura física y Arquitectura lógica.

Dentro de la arquitectura física podemos mencionar los diferentes tipos de redes:

- Computadoras personales (PCs)
- Redes de área local (LANs)
- Redes de área global (WANs)

Las Computadoras personales tienen las siguientes características

- Conexión esporádica a Internet.
- Necesaria la colaboración de otra computadora que esté conectada de forma continua a Internet (ISP Proveedor de Servicios Internet).
- Debe establecerse un diálogo con el ISP a través de algún canal de comunicación.

Las formas de conexión pueden ser:

1. Línea telefónica estándar.

- Necesario el uso de un hardware que actúe de interfaz entre la computadora y la línea telefónica (módem).
- Modems internos o externos.
- Velocidad baja

2. RDSI (Red Digital de Servicios Integrados).

- Línea digital.
- Sigue siendo necesario un hardware específico que actúe de interfaz entre la línea y la computadora.
- Hardware distinto del usado para la línea telefónica estándar (módem RDSI).
- Posibilidad de conexión a Internet y al mismo tiempo disponibilidad de línea telefónica.
- Posibilidad de videoconferencia punto a punto.

3. ADSL (Asymmetric Digital Subscriber Line).

- Línea de comunicación de alta velocidad.
- Utiliza como soporte el par de cables de cobre de la línea telefónica estándar.
- Velocidad de comunicación asimétrica

4. Cable.

- Utiliza uno de los canales de vídeo de un suministrador de acceso a cable.
- Suele suministrarse en unión de otros servicios (televisión, teléfono, etc.).

5. Satélite.

- Al igual que ADSL, se trata de un canal de comunicación asimétrico.
- Los datos de salida viajan generalmente a través de teléfono.
- La recepción se realiza a través de uno o más canales de vídeo emitidos por un satélite.

Redes de Área Local (LANs)

- Sirve para unir ordenadores situados en general en un mismo edificio o grupo de edificios.
- En la red pueden coexistir sistemas actuando como servidores y como clientes.
- Las computadoras se unen formando topologías (lineales, anillo, estrella, ...).
- Sigue siendo necesario un hardware dedicado que actúe como interfaz entre la computadora y la red (tarjeta de red).
- Velocidades típicas

1. Ethernet: 10 Mbaudios.

2. Token Ring: 4 y 16 Mbaudios.

3. Fast Ethernet: 100 Mbaudios.

4. Giga Ethernet: 1 Gbaudio.

- Nuevas tendencias: redes inalámbricas Redes de Área Global (WAN)

- Unión de múltiples LANs.

- La unión de múltiples WANs a nivel mundial es lo que se conoce como Internet.

Arquitectura lógica de Internet. Normalmente en Internet se usa la arquitectura cliente-servidor. Este tipo de organización se basa en que: entre todas las computadoras que están en la red, unos ofrecen servicios (los llamados servidores) y otros usan esos servicios (los denominados clientes). Cuando el cliente está viendo estas páginas, está accediendo a un servicio que le ofrece el servidor de páginas WEB. Los sistemas cliente-servidor están contruidos de tal modo que la base de datos puede residir en el servidor y ser compartida entre todos los usuarios.

La arquitectura de Internet a nivel lógico está estructurado por lo que se conoce como capas de servicio.

Modelo OSI de 7 capas:

1. Capa física: La Capa Física del modelo de referencia OSI es la que se encarga de las conexiones físicas de la computadora hacia la red, tanto en lo que se refiere al medio físico (medios guiados: cable coaxial, cable de par trenzado, fibra óptica y otros tipos de cables; medios no guiados: radio, infrarrojos, microondas, láser y otras redes inalámbricas);

2. Capa de datos: La capa de enlace de datos se ocupa del direccionamiento físico, de la topología de la red, del acceso a la red, de la notificación de errores, de la distribución ordenada de tramas y del control del flujo.

3. Capa de red: El cometido de la capa de red es hacer que los datos lleguen desde el origen al destino, aún cuando ambos no estén conectados directamente. Los dispositivos que facilitan tal tarea se denominan enrutadores, aunque es más frecuente encontrar el nombre inglés routers.

4. Capa de transporte: Su función básica es aceptar los datos enviados por las capas superiores, dividirlos en pequeñas partes si es necesario, y pasarlos a la capa de red, también se asegura que lleguen correctamente al otro lado de la comunicación.

5. Capa de Sesión: Esta capa es la que se encarga de mantener el enlace entre los dos computadores que estén transmitiendo archivos, ofrece varios servicios que son cruciales para la comunicación, como son:

- Control de la sesión a establecer entre el emisor y el receptor (quién transmite, quién escucha y seguimiento de ésta).
- Control de la concurrencia (que dos comunicaciones a la misma operación crítica no se efectúen al mismo tiempo).
- Mantener puntos de verificación, que sirven para que, ante una interrupción de transmisión por cualquier causa, la misma se pueda reanudar desde el último punto de verificación en lugar de repetirla desde el principio.

6. Capa de presentación: El objetivo de la capa de presentación es encargarse de la representación de la información, de manera que aunque distintos equipos puedan tener diferentes

representaciones internas de caracteres (ASCII, Unicode, EBCDIC), números (little-endian tipo Intel, big-endian tipo Motorola), sonido o imágenes, los datos lleguen de manera reconocible.

7. Capa de aplicación: Ofrece a las aplicaciones (de usuario o no) la posibilidad de acceder a los servicios de las demás capas y define los protocolos que utilizan las aplicaciones para intercambiar datos, como correo electrónico (POP y SMTP), gestores de bases de datos y servidor de archivos (FTP).

URL's

En la Web existe un sinnúmero de objetos, a los que se puede acceder mediante una variedad de protocolos. Para poder abstraer la idea de “objeto genérico”, se necesita disponer de conceptos que den sentido a nombres y direcciones en la Web. Un Identificador Universal de Recursos (URI) es un miembro de este conjunto universal de nombres, un Localizador Uniforme de Recursos (URL), es básicamente un caso particular de URI que expresa una dirección, mapeada a un algoritmo de recuperación del objeto que usa protocolos de comunicación a través de la red. Por último, un Nombre Uniforme de Recursos (URN) es algo que aún está en debate y que pretende definir un espacio de nombres para etiquetar objetos persistentes.

Un recurso es cualquier cosa distinguible. Por ejemplo, un documento electrónico, una imagen, un servicio, etc. Un recurso es, básicamente, un concepto que puede mapearse a una o a varias entidades, y que puede permanecer constante a pesar de ver alterados sus contenidos.

La razón por la que es necesario uniformizar identificadores, localizadores y nombres, es la gran cantidad de protocolos usados hoy en día para encontrar y recuperar recursos en la red. Los sistemas en los que se distribuye toda la información alcanzable utilizan una variedad de plataformas y formatos en constante evolución, que de no ser por los protocolos y convertidores de formato adecuados, no podrían ofrecer el acceso global que se da hoy en día. Sin embargo, esto no es posible de realizar al nivel de direcciones y nombres de los objetos, dado que son usados para identificarlos y distinguirlos. Además, estos nombres y direcciones son “transmitidos” de las más diversas maneras, desde memorándums hasta hipertextos, por lo que debe suponerse que estos identificadores tendrán una larga existencia.

De todas las ideas desarrolladas en este aspecto, aparece una característica común mapeable a la idea de un “objeto” y su correspondiente nombre/etiqueta/identificador. De esta manera se puede definir un conjunto de espacios de nombres en los que se dice que existen los objetos. URI: Identificadores Universales de Recursos. Corresponden a una forma de encapsular un nombre en un espacio de nombres registrados, y etiquetarlo con el espacio de nombres, produciendo un miembro del conjunto universal. La sintaxis universal permite el acceso a objetos disponibles utilizando protocolos existentes, y es capaz de ser extendida con el tiempo URL: Localizadores Uniformes de Recursos. Para los protocolos de acceso a Internet existentes, en la mayoría de los casos se hace necesario definir la codificación del algoritmo de acceso en algo suficientemente conciso para llamarse “dirección”. Las URIs que referencian objetos a los que se accede mediante protocolos existentes se conocen como URLs. URN: Nombres Uniformes de Recursos. Un URN difiere de una URL en que su propósito principal es etiquetar persistentemente un recurso con un identificador. Este identificador es obtenido de un conjunto de espacios de nombres definidos, cada uno de los cuales tiene establecida su propia estructura de nombres y procedimientos de asignación.

El esquema “urn” se encuentra reservado para establecer una estandarización de los requerimientos del espacio de nombres URN.

El formato de una URL básica es el siguiente:
Protocolo://host.dominio.dominiodelprimernivel/directorio/recurso

- **Protocolo.** Un protocolo es, básicamente, un lenguaje que utilizan las máquinas conectadas para comunicarse entre ellas. Con esto le indicamos a la máquina a la que nos conectamos el protocolo que debe usar. Dependiendo de este, probablemente tengamos que usar un programa distinto. Cada protocolo utiliza un puerto (una especie de “entrada” a la máquina) por default, y es a ese al que, en principio nos conectará al indicarle ese protocolo en la URL.
- **Host** La palabra Host hace referencia a cualquier clase de máquina conectada a la red y, por extensión, a su nombre.
- **Dominio** Se llama dominio al nombre que tiene en Internet una máquina o red de ellos. Una red puede contener a su vez a otras redes dentro de ella (esto es lo que hace que se llame a Internet “La red de redes”). Si este fuera el caso, podríamos encontrarnos con nombres como “subred.red.granred.com” que nos indicarían, de izquierda a derecha, qué red se encuentra dentro de la siguiente.
- **Dominio de primer nivel.** Todos los nombres de dominio, independientemente de si se refieren a una máquina o a una red de ellas, acaban con el llamado “dominio de primer nivel”. Este es el nombre que indica, en principio, cuál es la ocupación de la red (como son los casos, por ejemplo, de .com para empresas o actividades comerciales y .org para organizaciones) o a qué país pertenecen las autoridades que han otorgado ese nombre (por ejemplo, .es para España, .mx para México o .de para Alemania).
- **Directorio.** Se utiliza para localizar el archivo concreto, dentro de esa máquina, que queremos ver.
- **Recurso** se llama recurso a cualquier cosa que puedas encontrar en una máquina. Recurso es una página Web, una imagen o un programa para descargar. Pero también es un recurso el resultado de ejecutar un programa en una máquina remota o la imagen de una webcam. La mayoría de recursos que se pueden encontrar en la Web son, evidentemente, páginas.

Los protocolos utilizados más comunes son los siguientes:

- **FTP (File Transfer Protocol).** El protocolo FTP es uno de los más utilizados en Internet, ya que permite transferir archivos entre dos máquinas utilizando una conexión Internet. El protocolo dispone de una serie de formas estandarizadas por las que una máquina remota puede crear y cambiar directorios en la máquina local, así como transferir, copiar, mover y borrar Archivos. En realidad FTP utiliza comandos de texto plano para indicar las diversas órdenes, pero existen multitud de programas que simplifican su utilización mediante una interfaz gráfica en la que las operaciones de transferencia se reducen a procesos de cortar y pegar y en las que recorrer el árbol de directorios de la máquina remota es tan fácil como utilizar el navegador en la propia máquina.
- **SMTP (Simple Mail Transfer Protocol).** Protocolo encargado de transmitir los mensajes de correo de un servidor a otro. Una vez que los mensajes están en el servidor de destino se utiliza otro protocolo (POP) para llevarlo a la máquina del cliente final.
- **POP (Post Office Protocol).** Es el protocolo que permite a un usuario leer el correo que llega a su servidor. Es un protocolo para las comunicaciones Servidor a Usuario. Cuando recibimos un e-mail queda almacenado en el servidor hasta que conectamos con él mediante el cliente de correo y nos autentificamos (proporcionamos un nombre de usuario y contraseña correctos). Después de esto POP es utilizado para transferir los datos desde el servidor al buzón de correo entrante de nuestra

propia máquina. Eventualmente una vez recibida la copia es posible ordenar al servidor que borre los Archivos originales.

- IRC (Internet Relay Chat). Es un protocolo de comunicación en tiempo real basado en texto, que permite debates en grupo o entre dos personas y que está clasificado dentro de la Mensajería instantánea. Es un sistema de charlas ampliamente utilizado por personas de todo el mundo. Los usuarios del IRC utilizan una aplicación cliente para conectarse con un servidor, en el que funciona una aplicación IRCD (IRC Daemon o servidor de IRC) que gestiona los canales y las conversaciones.
- Telnet. Es un servicio que permite registrarse en un servidor y acceder a sus diferentes recursos. Los investigadores aprecian mucho este servicio cuando buscan información sobre librerías y archivos. Telnet es especialmente importante en la conexión con información almacenada en macrocomputadoras. El servicio de Telnet tiene su propio protocolo llamado: Telnet. No es posible buscar un servidor de Telnet a través de los proveedores de servicios en línea. Para buscar a un servidor de Telnet, se debe establecer una sesión.
- WWW. El World Wide Web (WWW) fue creado por científicos del CERN quienes necesitan compartir y acceder información sobre investigaciones a través de una interfaz común. Al usar una interfaz común, los investigadores simplificaban los pasos necesarios para acceder a los diferentes servicios disponibles en Internet. Los programas de aplicación raramente referencian a los hosts, buzones de correo y otros recursos por su dirección de red en formato binario. En vez de esto utilizan cadenas de caracteres. A pesar de esto la red por si misma solo entiende direcciones binarias, por lo que se requiere de algún mecanismo para convertir las cadenas de caracteres en direcciones de red. En sus inicios ARPANET, esto se realizaba mediante un simple Archivo, hosts.txt, que listaba todos los hosts y sus respectivas direcciones IP. Como se puede suponer, cada día se tenía que estar verificando y actualizando este Archivo. Además para un número razonable de máquinas, esta solución podía ser viable, pero el aumento de estaciones conectadas a la red aumentaba considerablemente cada día, por lo que el tamaño del Archivo podía llegar a ser muy grande. Para solventar estos problemas se invento el DNS (Domain Name System también conocido como Domain Name Server o Domain Name Service).

La esencia del DNS se basa en la invención de un esquema jerárquico, asignando nombres que se basan en dominios y en un sistema de base de datos distribuido para implementarlo. Su principal función consiste en mapear nombres de hosts y direcciones e-mail en direcciones IP, aunque se le puede dar otras muchas utilidades. De una forma breve, su manera de trabajar es la siguiente. Para mapear un nombre en una dirección IP, un programa llama a un procedimiento pasándole el nombre como parámetro. Este procedimiento entonces envía un paquete UDP (User Datagram Protocol) a su servidor DNS local, que busca el nombre y retorna la dirección IP al procedimiento, que se lo devuelve a la aplicación. Una vez que el programa de aplicación tiene la dirección IP puede establecer una conexión TCP con el destino. Conceptualmente, Internet esta dividida en numerosos dominios, que cada uno cubre cientos de hosts, a su vez cada dominio esta dividido en subdominios, y estos a su vez subdivididos de forma recursiva.

Existen dos tipos de dominios superiores, genéricos y territoriales.

Los dominios territoriales son los dominios mantenidos por cada país. Estos dominios territoriales son utilizados por las organizaciones y empresas que desean establecerse en Internet o que quieren proteger la identidad de su marca o su nombre comercial en un país concreto. Los dominios están definidos en ISO 3166. Algunos ejemplos de dominios territoriales son los

siguientes:

es : España

mx : México

ar : Argentina

uk : Reino Unido

at : Austria

il : Israel

ch: Suiza

Los dominios genéricos son los dominios básicos en Internet. Cualquier empresa que realice negocios a través de Internet o que tenga intención de hacerlo debe registrarse en uno de estos dominios genéricos, dada la popularidad que han adquirido. Los dominios genéricos son:

com : Comercial

edu : Educational institutions

gov : U.S. Federal Government

int : International Organizations

mil : U.S. Armed Forces

net : Network Provider

org : Nonprofit Organizations

Cada nombre de un dominio puede tener hasta 63 caracteres, y el nombre del dominio completo no puede sobrepasar los 255 caracteres. En la teoría, un solo servidor puede tener la base de datos DNS completa, y responder a todas las consultas. En la práctica, este servidor podría estar sobrecargado o infrautilizado. Además si este servidor cayera, toda Internet quedaría colapsada.

Para solucionar el problema asociado a tener una sola fuente de información, el espacio de nombres DNS se divide en zonas. De forma que cada zona tiene un servidor principal con la información en su disco, y uno o varios servidores secundarios, que consiguen la información del servidor principal. Para una mejor eficiencia, algunos servidores de una zona pueden ser colocados en otra.

El Network Information Center – México, (NIC-México) es la organización encargada de la administración del nombre de dominio territorial (ccTLD, country code Top Level Domain) .MX, el código de dos letras asignado a cada país según el ISO 3166. Entre sus funciones están el proveer los servicios de información y registro para .MX así como la asignación de direcciones de IP y el mantenimiento de las bases de datos respectivas a cada recurso.

Un ejemplo en que una máquina accesa una URL es el siguiente:

Un cliente pregunta por un dominio .MX (www.negocio.com.mx), esta respuesta tiene que ser contestada por el servidor de nombres de la red a la que pertenece el cliente que pregunta. Como primera instancia, el servidor de la red local, deberá preguntar a los servidores raíz por el dominio que está buscando (negocio.com.mx). Estos servidores conocen la información de los diferentes NICs en el mundo, así como de los dominios genéricos, por lo que le indican al servidor de la red local, dónde puede encontrar más información de los dominios bajo .MX y lo dirigen con los servidores de NIC México. Los servidores de NIC México son también los servidores para el .com.mx, por lo que buscan en su base de datos la información relacionada al dominio negocio.com.mx. En la información que se recaba, se tiene que el dominio negocio.com.mx tiene como servidores registrados ns1.negocio.com.mx y ns2.negocio.com.mx, estos servidores conocen la información referente al dominio negocio.com.mx, por lo que dan la respuesta a la pregunta

de dónde encontrar <http://www.negocio.com.mx>. Cabe mencionar que quien realizó todo el proceso de la búsqueda fue el servidor de la red local. Ya con la respuesta de dónde localizar a <http://www.negocio.com.mx>, el servidor le proporciona esta respuesta al cliente, con esta información, el cliente, puede tener una comunicación directa con el host para solicitar algún servicio, que en este caso, es el servicio de web. Con esto, se termina el proceso de resolución de un nombre de dominio.

Métodos http.

Los métodos que utiliza el protocolo http son los siguientes:

1. GET: se utiliza para recuperar el contenido de un recurso estático (página HTML, imagen, vídeo, etc), o bien ejecutar un programa o script en el servidor (como programas CGI, páginas ASP, JSP, PHP, etc.). Es la operación más común, que se lanza cada vez que pulsamos en un enlace y navegamos a una página web.
2. HEAD: es una operación especial que tan sólo nos recupera información del recurso, como el tamaño, la fecha de modificación, tipo, etc. Lo suelen utilizar los navegadores o servidores para comprobar el estado de su caché u otras operaciones.
3. POST: envía información desde el cliente al servidor web, como pueden ser los datos de un formulario.
4. PUT: almacena recursos en el servidor.
5. DELETE: borra recursos del servidor.

Uno de los problemas clásicos en el desarrollo de web sites y aplicaciones web es la pérdida de persistencia cuando el usuario pasa de una página a otra. Debido a las características de diseño del protocolo HTTP que fuerza una nueva conexión y desconexión por cada request no es posible saber quien está accediendo a que página o en que lugar esta cada usuario del sitio. Mantener persistencia a lo largo de la navegación del sitio ha sido uno de los temas más complejos e importantes en el desarrollo de aplicaciones web.

Uno de los mecanismos más usados para mantener persistencia es el mecanismo de cookies, inventado por Netscape y hoy en día aceptado por casi todos los browsers, en especial los más populares. El concepto es que mediante un header del protocolo HTTP el server pueda almacenar información en el cliente. A esta información que el server guarda en el cliente se la denomina "cookie". Las cookies pueden habilitarse o deshabilitarse desde el browser por lo que algunos usuarios no lo soportan, son de uso bastante general en muchos sitios web a punto tal que en algunos sitios si el usuario no tiene habilitadas las cookies prácticamente no puede utilizar la mayoría de los servicios del sitio. Cuando el server envía un header con un cookie el browser, si acepta cookies, guarda la información enviada en un archivo de texto con un formato especial.

Cada vez que el browser solicita una página del dominio que envió la cookie reenvía la cookie al sitio, de esta forma es posible mantener persistencia. La información que puede guardarse en una cookie esta limitada por lo que habitualmente se utiliza la misma para mantener el identificador de sesión

del usuario almacenándose el resto de los datos necesarios en el servidor usando la session id de la cookie como clave.

Los usos más frecuentes de las cookies son:

- Llevar el control de usuarios: cuando un usuario introduce su nombre de usuario y contraseña, se almacena una cookie para que no tenga que estar introduciéndolas para cada página del servidor. Sin embargo una cookie no identifica a una persona, sino a una combinación de computadora y navegador.
- Ofrecer opciones de diseño (colores, fondos, etc) o de contenidos al visitante.
- Conseguir información sobre los hábitos de navegación del usuario, e intentos de [spyware](#), por parte de agencias de publicidad y otros.

Cuando las cookies han caducado, estas no son enviadas al navegador; por lo tanto, la caducidad de las cookies puede ser pensada como un límite de tiempo en el que una de ellas puede ser usada. La cookie puede luego ser renovada después de que este límite haya pasado. Algunos sitios prefieren que las cookies caduquen en tiempos más cortos por razones de seguridad. Las cookies no se envían al navegador si ellas están bajo estas condiciones:

- Al finalizar una sesión de usuario: por ejemplo, cuando se cierra el navegador (si esta no es persistente)
- Se ha fijado una fecha de caducidad y esta ha pasado.
- La fecha de caducidad es cambiada a una fecha anterior (por el servidor)
- Esta se borra por orden del usuario.

Conclusión

En la actualidad nos damos cuenta de que hablar de tecnologías web engloba muchos campos y que sus áreas son diversas y complejas, desde solo consultar información con un navegador ya sea Firefox, safari, opera, chrome, buscar una imagen, consultar sitios de empresas, hacer compras electrónicas, jugar, conocer lugares en 3D, así como consultar datos estadísticos, bases de datos, sitios de gobierno, etc., etc... Las tecnologías juegan un papel muy importante en la web, debido a estas herramientas podemos acceder al internet, que no es solo sentarte frente a un ordenador y abrir el internet sino que hay un gran trabajo detrás para hacer posible este tipo de tecnología, donde miles de personas trabajan actualizando datos, creando los métodos para mantener todos los enlaces y las comunicaciones.

Referencias

- Manuel Montenegro.
- (2018).
- Introducción a las tecnologías Web.
- 26/06/2018,
- Instituto Tecnológico de Matehuala Sitio web:
<https://programacionwebisc.wordpress.com/1-2-2-urls/>

