



01 Funcionamiento de Internet

Introducción a los protocolos:
HTTP y DNS

Definición de una web

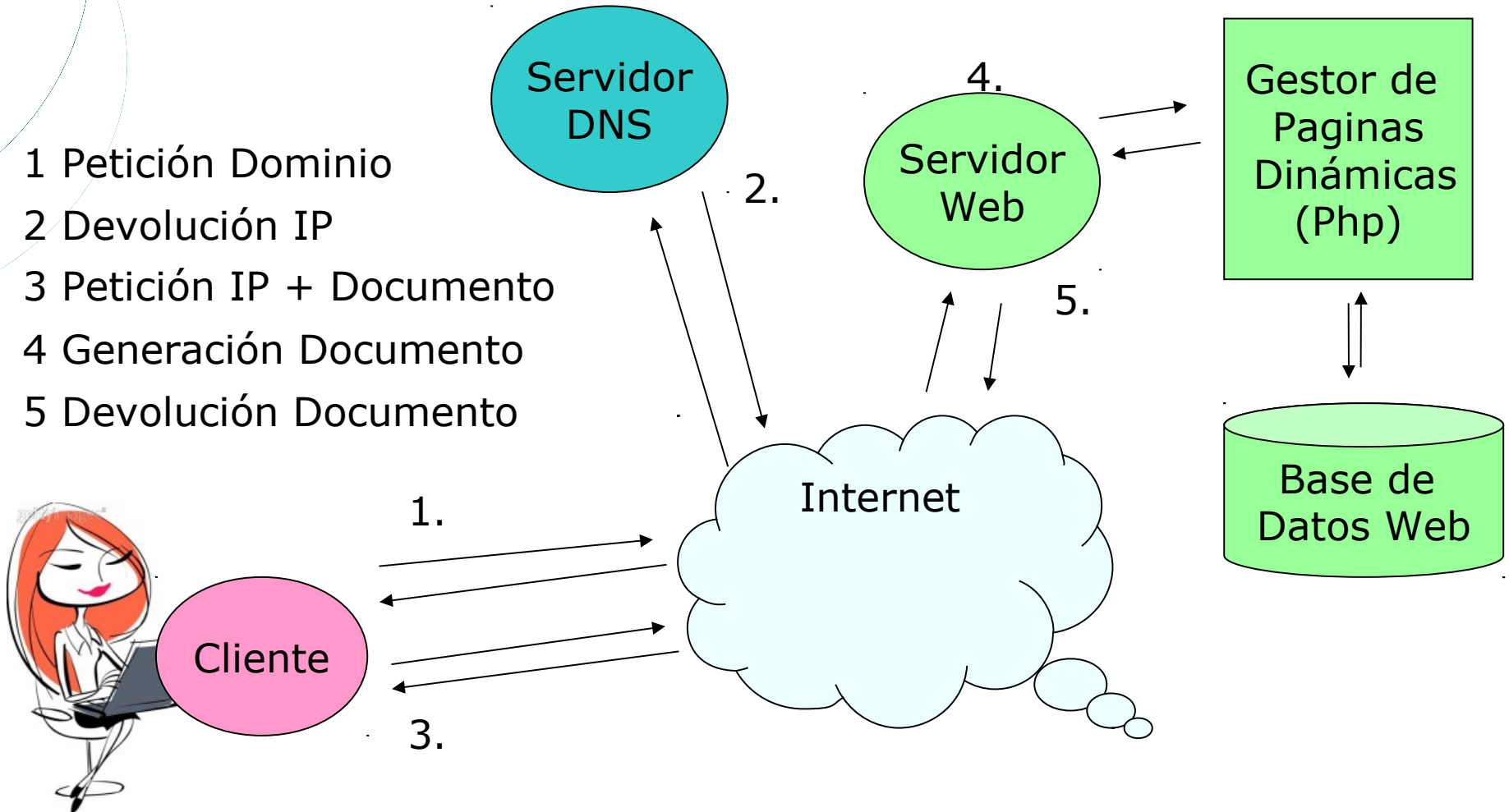
- Una web no es mas que un conjunto de documentos HTML y otros incrustados en un servidor http (como apache), que son enviados usando el protocolo http, a los usuarios que lo solicitan mediante un software cliente que los interpreta llamado Navegador.

Definición de una web

- La forma de comunicarse entre el ordenador cliente y el servidor es mediante la dirección IP de ambos, de cara al usuario un nombre que se resuelve usando el protocolo DNS

Funcionamiento de un Acceso Web

- 1 Petición Dominio
- 2 Devolución IP
- 3 Petición IP + Documento
- 4 Generación Documento
- 5 Devolución Documento



El protocolo DNS

La guía telefónica de internet

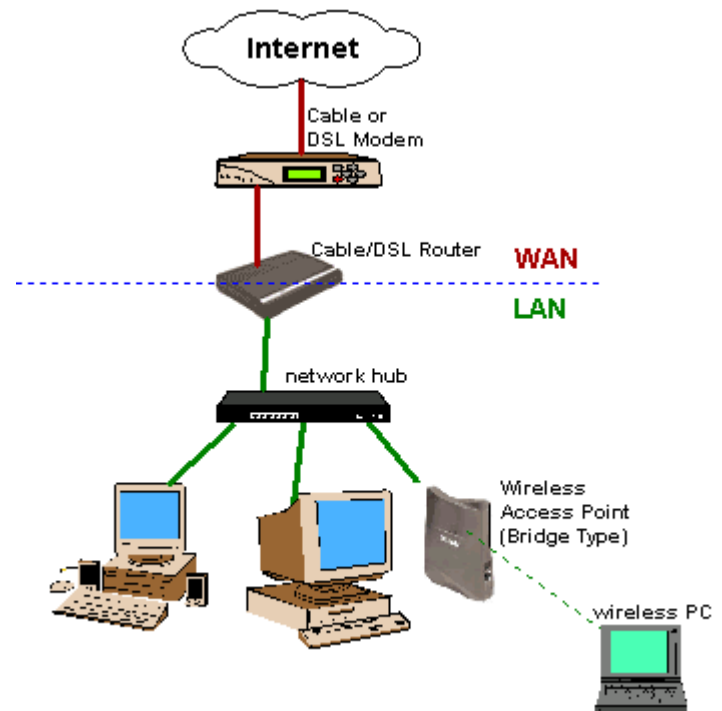


Direccionamiento en Internet

- Todo Internet funciona con el protocolo IP, que establece que las máquinas están identificadas por direcciones IP, que constan de 32 bits, (cuatro números decimales separados por puntos) en IPV4 en IPV6 son 128 bits.
- Ejemplos:
 - 127.0.0.1
 - 234.34.54.3
 - 2001:0DB8:0:0:0:0:1428:57ab

Direccionamiento en Internet

En realidad no hay una dirección IP por cada máquina de Internet, sino una (o más) por cada "punto de acceso a red", o interfaz de red



Problemática

Direccionamiento en Internet

- Direcciones IP son Muy engorrosos para recordar por los usuarios
 - 127.0.0.1
 - 234.34.54.3
 - 2001:0DB8:0:0:0:0:1428:57ab
- Solución: DNS

ICANN



- **ICANN** (*Internet Corporation for Assigned Names and Numbers*) es una organización sin fines de lucro a la cual el gobierno de los EEUU le encargó la tarea de coordinar ciertas funciones técnicas de internet, entre ellas, la gestión de **DNS**.
- Sólo empresas acreditadas por la ICANN están autorizados para registrar ciertos dominios
- Hasta **1999**, una de ellas, **Network Solutions Inc** era una especie de monopolio mundial en la gestión de nombres de dominio internacionales del primer nivel pero, a partir de esa fecha, se liberó el mercado entrando en juego nuevas empresas con la consiguiente baja en los precios

ICANN

Network**Solutions**.

DNS

- El sistema y protocolo DNS (Domain Name System) es una base de datos que asocia las direcciones IP en Internet con nombres textuales fáciles de recordar para las personas.
- Cuando un usuario quiera referirse a una máquina de Internet, usa su nombre, e internamente el sistema DNS se encargaría de averiguar la dirección IP.

Funciones adicionales al DNS

- Averiguar el nombre de una máquina a partir de su dirección IP.
- Buscar el servidor de correo
 - Cuando enviamos un correo electrónico debemos averiguar cuál es la dirección del servidor de correo del destinatario a partir de la dirección de correo del destinatario (ramon@ola.com)

Problema DNS

- En el principio de Internet, DNS se regia mediante una tabla de equivalencias en varios servidores.
- Piense ahora por un momento el tamaño que puede tener esta base de datos llamada DNS, encargada de albergar las equivalencias de nombres de máquina a dirección IP.

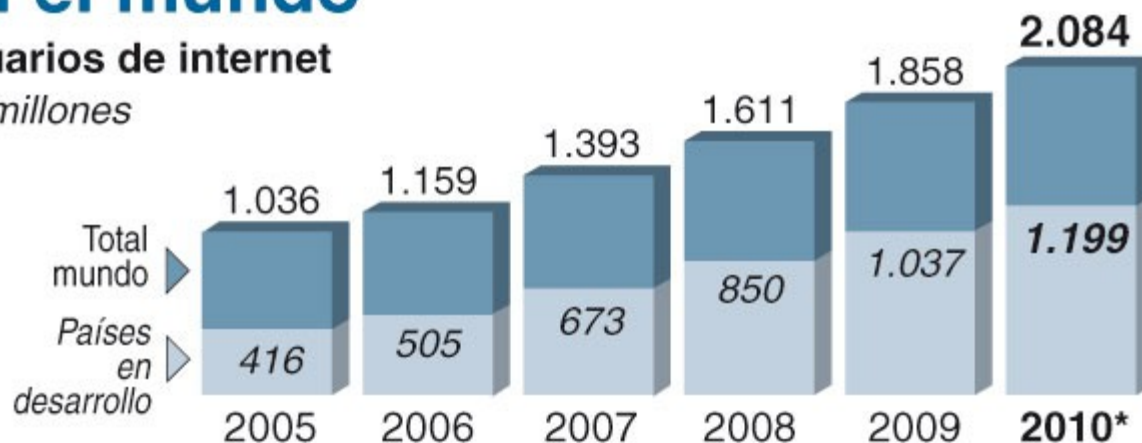
Problema DNS

- Teniendo en cuenta que Internet cuenta con varios cientos de millones de máquinas puede suponer que dicha base de datos es inmensa.
- Pero el problema es el volumen de peticiones que recibe a lo largo del día. Por cada página web que visita se generan posiblemente varias consultas al DNS, así como para cada correo enviado, o cualquier otra actividad en Internet.

2.000 millones de internautas en el mundo

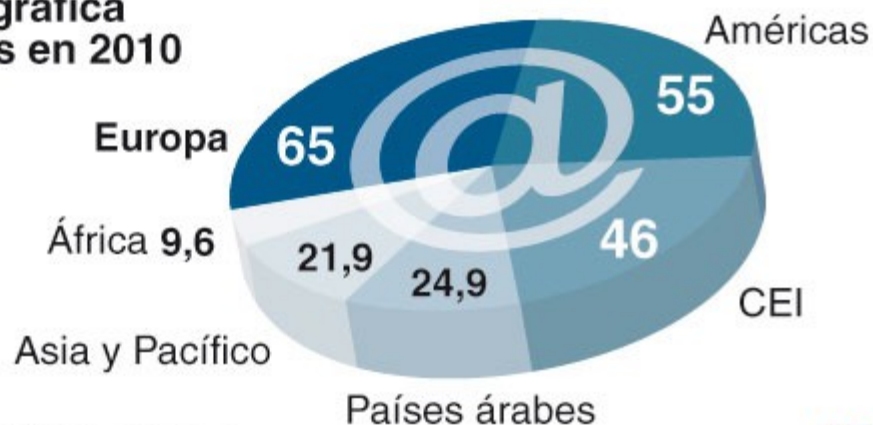
Usuarios de internet

En millones



Distribución geográfica de los internautas en 2010

Cantidad de usuarios de internet por 100 habitantes*



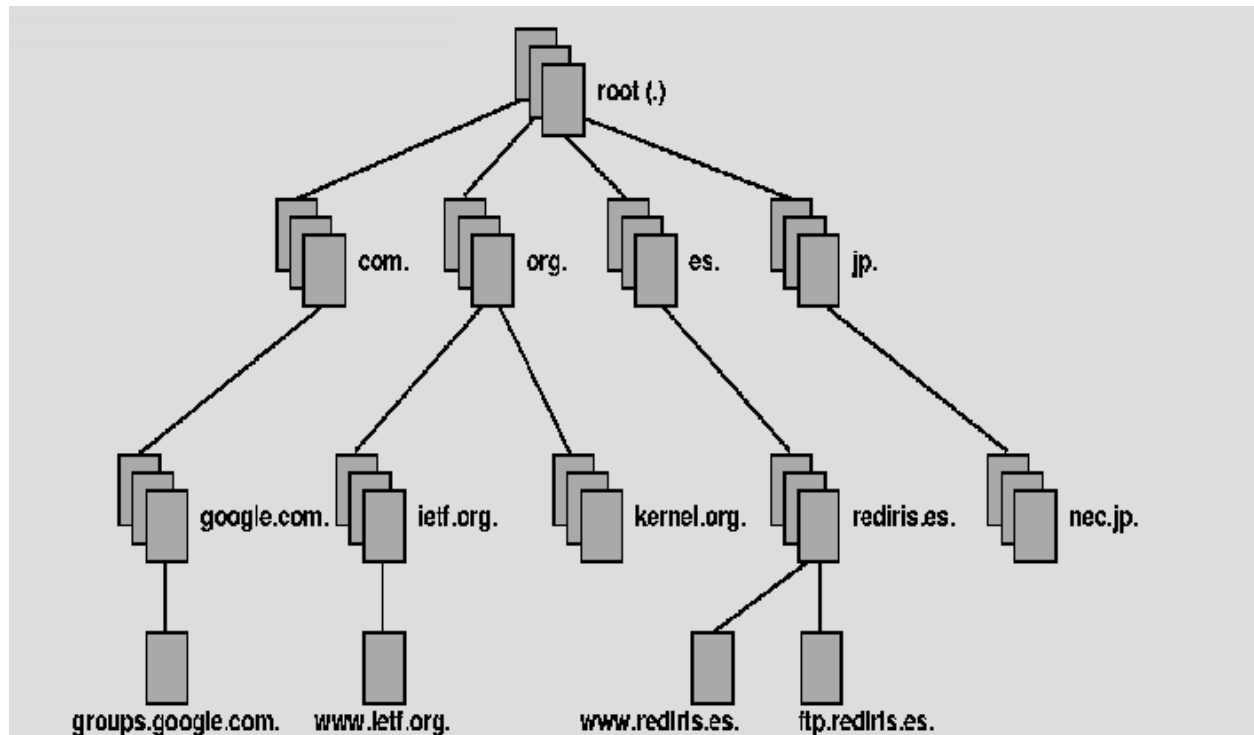
*estimaciones

Fuente: Unión Internacional de Telecomunicaciones

AFP

Solución DNS

- La solución es diseñar el sistema usando una Base de datos jerárquica, descentralizada y replicada



Solución DNS

- Existen servidores Dns de diferente nivel. Y cada nivel solo se debe preocupar de conocer las dirección de los servidores Dns inferiores.
- Por ejemplo el servidor .com conocería donde se encuentran flores.com y este a su vez conoceria la maquina que resuelve smtp.flores.com

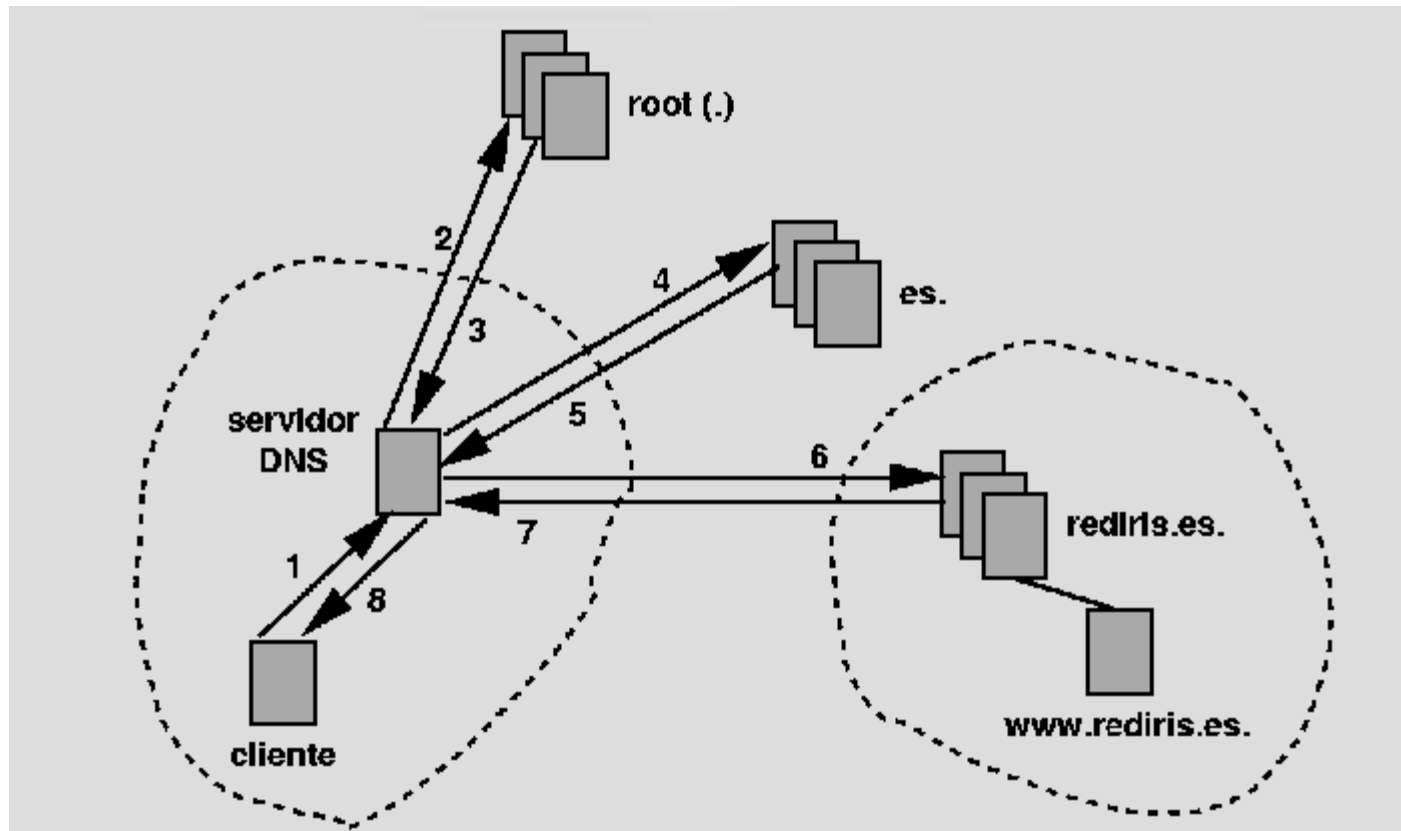
Solución DNS

- En el tope de la jerarquía está la raíz, representada por un punto. El servidor de nombres raíz sólo tendrá información acerca de cómo llegar a los servidores de nivel inmediatamente inferior.
- Los conocidos como dominios de primer nivel, o TLD (Top Level Domain). Existen más de doscientos TLD definidos actualmente, entre **dominios nacionales** (uno por cada nación, definidos por su identificador de dos letras según el estándar ISO-3166) y **dominios genéricos** (com, net, org, gov, mil, edu, y más).
- Cada uno de estos dominios está gestionado por una empresa u organismo diferente, y configurado en uno o más servidores DNS repartidos por Internet.

Solución DNS

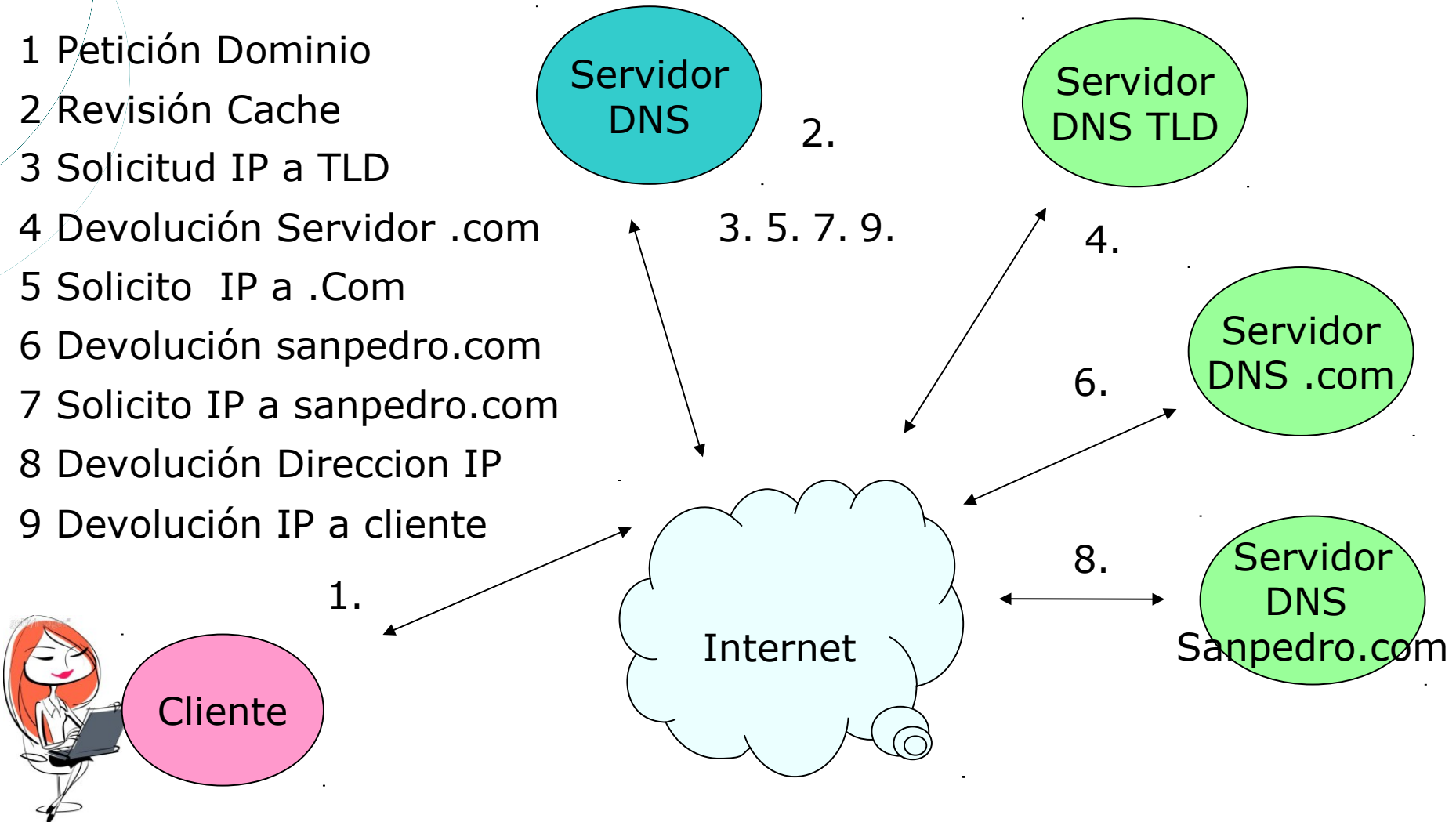
- Los agencias que gestionan los dominios de primer nivel o TLD pueden delegar parte de su espacio de nombres a empresas para que las gestionen ellas mismas y así ahorrar recursos.
- Cuando un servidor esta capacitado para dar la información de un dominio se dice que es authoritative.

Ejemplo de llamada DNS



Funcionamiento de una Consulta DNS

- 1 Petición Dominio
- 2 Revisión Cache
- 3 Solicitud IP a TLD
- 4 Devolución Servidor .com
- 5 Solicito IP a .Com
- 6 Devolución sanpedro.com
- 7 Solicito IP a sanpedro.com
- 8 Devolución Direccion IP
- 9 Devolución IP a cliente





DNS Inverso

Práctica

- Usar la herramienta nslookup para ver la dirección IP de los dominios siguientes.
 - www.google.es
 - www.google.com
 - www.yahoo.es
 - www.terra.es
 - www.aol.com
 - www.hotmail.com

El protocolo HTTP

El cartero de Internet



El protocolo http

- Una vez obtenemos la dirección IP de la máquina que deseamos el documento Utilizamos el protocolo HTTP para solicitarlo.

HTTP, *HyperText Transfer Protocol*

- El protocolo de transferencia de hipertexto (HTTP, HyperText Transfer Protocol) es el protocolo usado en cada transacción de la Web (WWW).
- HTTP fue desarrollado por el consorcio W3C y la IETF, colaboración que culminó en 1999 con la publicación de una serie de RFC, siendo el más importante el RFC 2616, que especifica la versión 1.1.

El protocolo HTTP

- El protocolo HTTP es el usado en la transferencia de paginas web entre el cliente y el servidor.
- Se usa tanto para que el navegador pida una página a un servidor, como para que éste envíe la página solicitada al navegador..

El protocolo HTTP

- Está basado en el envío de comandos y respuestas en texto ASCII.
- Si bien cada tipo de contenido enviado (archivo en formato HTML, imágenes, documentos PDF) se enviará tal cual está almacenado en el servidor.
- Es decir, los archivos binarios se envían tal cual, aunque los comandos y las respuestas del protocolo HTTP vayan siempre en texto.

El protocolo HTTP

- Cuando un navegador conecta con un servidor web remoto le envía un fichero de texto donde ...
 - le solicita la página cuya URL el usuario ha tecleado (o la indicada en el enlace seleccionado con el ratón).
 - le informa de detalles del navegador, de la página que solicita.

Petición HTTP típica

- GET / HTTP/1.1
- Host: www.google.com
- User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US;
- Accept:text/xml,application/xml,application/xhtml+xml,text/html;q=0.9, text/plain;q=0.8,video/x-mng,image/png,image/jpeg,image/gif;q=0.2, text/css,*/*;q=0.1
- Accept-Language: es-es, en-us;q=0.66, en;q=0.33
- Accept-Encoding: gzip, deflate, compress;q=0.9
- Accept-Charset: ISO-8859-15, utf-8;q=0.66, *;q=0.66
- Keep-Alive: 300
- Connection: keep-alive

Peticion Inicial

- Se informa del navegador en uso (**User-Agent**)
- En qué idioma se prefiere visualizar la página que se está solicitando (**Accept-Language**)
- Las cabeceras **Accept-Encoding** y **Accept-Charset** y **Accept** le indican al servidor qué tipos de contenidos el navegador entiende y sabe representar.

Peticion Inicial

- La pareja de cabeceras **Keep-Alive y Connection**, permite recibir muchos recursos (páginas HTML, imágenes, etc.) a través de la conexión TCP recién establecida.
- En la versión 1.0 del protocolo HTTP era necesario establecer y terminar una conexión TCP por cada recurso, lo que resultaba tremendamente ineficiente.

Contestación Servidor

- El servidor recibe todas estas cabeceras, las interpreta, y devolverá al navegador una respuesta estándar HTTP (indicando si la petición tuvo éxito o no), y a continuación la página solicitada.
- En el caso de la página principal implícita (/), el servidor tendrá configurada un archivo, index.html, que enviará en caso de no especificarse una página concreta.

Contestación Servidor

HTTP/1.1 200 OK

Date: Sun, 10 Nov 2002 22:50:55 GMT

Server: Apache/1.3.26 (Unix) mod_bwlimited/1.0

PHP/4.2.2 mod_log_bytes/0.3

FrontPage/5.0.2.2510 mod_ssl/2.8.9

OpenSSL/0.9.6b

Content-Type: text/html

Age: 130

Connection: close

<-- archivo index.html que contiene la página principal del sitio -->

Códigos http ok 2XX

- 200 OK
 - Respuesta estándar para peticiones correctas.
- 201 Creado
 - La petición ha sido completada y ha resultado en la creación de un nuevo recurso.
- 202 Aceptada
 - La petición ha sido aceptada para procesamiento, pero este no ha sido completado.
- 203 Información no autoritativa (desde HTTP/1.1)
- 204 Sin contenido
- 205 Recargar contenido

Códigos http redirección 3XX

- 301 Moved Permanently

Códigos http error cliente 4XX

- 400 Solicitud incorrecta
- 401 No autorizado
 - Similar al 403 Forbidden, pero específicamente para su uso cuando la autenticación es posible pero ha fallado o aún no ha sido provista.
- 402 Pago requerido
 - Proyecto fallado de Esquema de Dinero electrónico. este código nunca se utilizó.
- 403 Prohibido
 - La solicitud fué legal, pero el servidor se rehusa a responderla.
- 404 No encontrado
 - Recurso no encontrado.

Códigos http error servidor 5XX

- 500 Error interno
 - Es un código comúnmente emitido por aplicaciones empotradas en servidores web, mismas que generan contenido dinámicamente, por ejemplo aplicaciones montadas en IIS o Tomcat, cuando se encuentran con situaciones de error ajenas a la naturaleza del servidor web.
- 501 No implementado
- 503 Servicio no disponible

Práctica

- Utilizar el plugin de Firefox, LiveHttpHeaders para conseguir las cabeceras HTTP. <http://livehttpheaders.mozdev.org/>



Práctica

- Busca webs que te den los diferentes tipos de error y ok y haz un copiado a un Documento.
- Prueba
 - www.google.com
 - es.wikipedia.org/wiki/
 - www.iessalduba.com/ESI/
 - Adjuntar un archivo a hotmail
 - www.deeptraining.com/images/

