

“A la avenida Google, por favor. Y rápido”

Miguel Jaque Barbero – diciembre 2008

Internet se ha convertido en un elemento más de la vida diaria. Pero apenas sabemos nada de su funcionamiento. ¿Cómo encuentra nuestro navegador las páginas que le solicitamos entre los miles de millones de páginas de Internet?



Cuando escribimos una dirección web en nuestro navegador, no nos sorprende que en pocos instantes aparezca la página solicitada en nuestra pantalla. Pero en Internet hay más de 500 millones de servidores web, cada uno de ellos con miles o incluso millones de documentos. ¿Cómo puede nuestro navegador encontrar el documento que buscamos? ¿Y cómo puede hacerlo tan rápido?

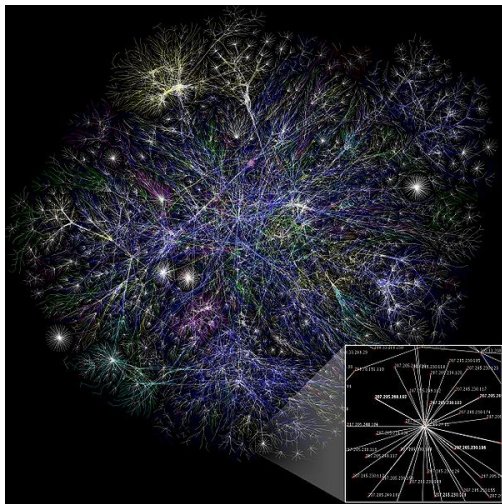


Ilustración 1: Mapa de servidores en Internet

Nadie sabe exactamente cuántas páginas hay hoy en Internet, el número crece cada día. Pero lo cierto es que nuestros navegadores son capaces de presentarnos el documento que buscamos en cuestión de segundos. Y lo consiguen utilizando unos protocolos de comunicación relativamente simples, pero muy potentes. Veamos cómo se las apaña nuestro navegador para encontrar lo que buscamos.

Todo empieza cuando escribimos la dirección web (por ejemplo www.google.es) en el navegador y pulsamos “Enter”. La dirección web que

escribimos se denomina técnicamente URL (*Uniform Resource Locator*) e identifica a cada una de las páginas y documentos que hay en Internet. Encontramos direcciones web en las páginas de resultados de un buscador, como enlaces en otras páginas o en la publicidad de cualquier empresa.

El proceso para localizar la página que solicitamos tiene dos pasos. En primer lugar, hay que traducir la dirección web, compuesta por letras y palabras que los humanos podemos recordar, por una dirección totalmente numérica que es lo que los ordenadores pueden manejar. Esta dirección numérica se denomina técnicamente IP (*Internet Protocol*). La dirección numérica es única para cada ordenador conectado a Internet. Por ejemplo, la dirección numérica de Google es 66.249.89.99. Se puede comprobar escribiendo los números, con los puntos que los separan, en el navegador. Veremos la misma página que si escribimos www.google.es. Y una vez obtenida la dirección numérica, el segundo paso será enviar al ordenador que tiene la página que buscamos un mensaje solicitándosela.

¿Cómo se traduce la dirección web en una dirección numérica?

A este proceso de traducción se le denomina “resolver el nombre”. Obviamente nuestro ordenador no puede conocer todas las direcciones numéricas de todos los ordenadores de Internet. No habría memoria suficiente para ello.

Aunque sí puede acordarse de las últimas direcciones web que hemos visitado y, si este es el caso, tendrá ya memorizada la dirección numérica que les corresponde. Asunto resuelto.

Pero, si no hemos visitado recientemente esa dirección web, nuestro navegador tendrá más trabajo. Para ello, todos los ordenadores conectados a Internet tienen registradas dos o más direcciones numéricas que corresponden a “*Servidores de Nombres*”.

Los *Servidores de Nombres*, técnicamente llamados DNS (*Domain Name Server*), son ordenadores que se encargan de traducir las direcciones web en direcciones numéricas. Pero ¿cómo lo hacen?

Los Servidores de Nombres tienen un directorio con las direcciones web que les han preguntado y su dirección numérica. Y como, lo más probable, es que la dirección ya se la hayan preguntado alguna vez, ya estará en su directorio. Asunto resuelto.

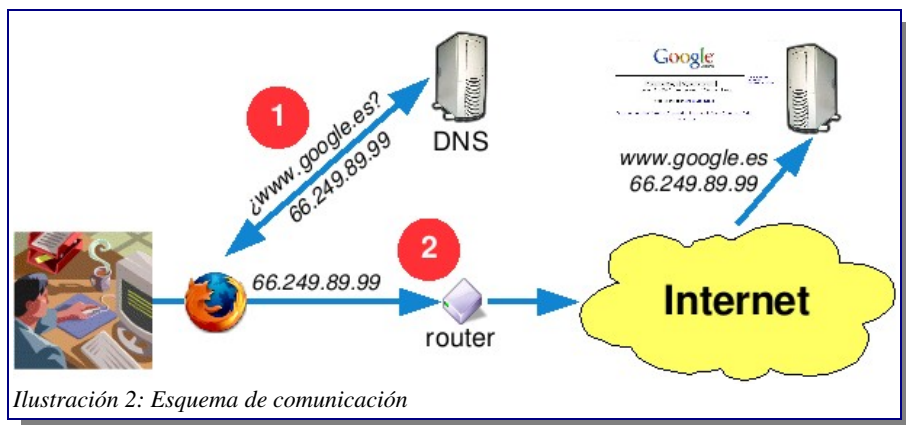
¿Y si la dirección no está en la memoria del Servidor de Nombres?

Si el Servidor de Nombres al que preguntamos no tiene esa información, tendrá que preguntar a otros Servidores. Para ello, se quedará con las dos últimas palabras antes de la barra, que forman el “*dominio*”. Por ejemplo, la dirección web para el artículo de Albert Einstein en la Wikipedia es es.wikipedia.org/wiki/Einstein. En este caso, el *dominio* es “*wikipedia.org*”. Cada dominio tiene Servidor de Nombres propio que es quien

conoce las direcciones numéricas de todos los ordenadores de ese dominio. Dominios como *Google.com* o *Microsoft.com* pueden tener varios cientos de ordenadores conectados a Internet.

Este trabajo corresponde al *router*. El *router* es el elemento que conecta nuestra red local, en la que está nuestro ordenador, con otras redes. Internet es un inmenso conjunto de redes interconectadas, y los elementos que

de Internet intercambian entre ellos esta información y crean sus “*tablas de enrutamiento*”. Si el router B sabe que tiene conexión con los ordenadores B1, B2 y B3, al enviar esta información a A, éste la añadirá a su tabla de *enrutamiento*, en la que figurará que A puede llegar a los ordenadores B1, B2 y B3 enviando los mensajes a través del router B.



Aun así, podría ocurrir que nuestro Servidor de Nombres no conociera al Servidor del dominio que buscamos. Actualmente hay más de 150 millones de dominios registrados y el número crece en un 30% cada año. Pero no tampoco hay problema, si no sabe a quien preguntar, lo consultará con un *DNS raíz*. Estos son Servidores de Nombres que saben qué Servidor de Nombres corresponde a cualquier dominio. Actualmente hay 13 *DNS raíz* que son conocidos por todos los Servidores de Nombres de Internet. Y a pesar de que responden miles de consultas por segundo, trabajan bastante bien gracias a que el resto de Servidores de la red recuerdan las respuestas obtenidas y las comparten con otros Servidores.

Al final, bien porque la recuerda de consultas anteriores o bien porque la consulta a un Servidor de Nombres, nuestro navegador obtiene la dirección numérica asociada al ordenador que alberga el documento que buscamos.

Y con la dirección numérica ¿cómo encontramos el ordenador que buscamos?

En Internet hay millones de ordenadores, cada uno con su propia dirección numérica. ¿Cómo utilizamos esta dirección para encontrar el ordenador al que corresponde?

conectan unas redes con otras son los *routers*. Un *router* no es más que un ordenador que conecta varias redes entre sí, y que está especializado en llevar información hasta su destino.

Pues una vez identificada la dirección numérica del destino, nuestro navegador se limitará a enviar un mensaje a esa dirección pidiéndole que le envíe la página que queremos visualizar. Esta petición le llegará a nuestro *router*. Ahora el *router* tiene que encontrar el camino, entre los miles de millones de ordenadores de Internet para entregar el mensaje al ordenador de destino. A este trabajo de encontrar el camino se le llama “*enrutamiento*”.

Cada uno de los *routers* de Internet conoce las conexiones que tiene con otros ordenadores (*routers*, servidores y clientes). Por ejemplo, el router A sabe que tiene conexiones con los ordenadores B, C, y D. Cada cierto tiempo los *routers*

Pero todavía faltan dos detalles para terminar. Internet es un conjunto de redes muy interconectadas, por lo que casi siempre existen varios caminos para llegar a un mismo destino. Los *routers* deciden cual es el mejor camino añadiendo a su *tabla de enrutamiento* la información del “*coste*” del enlace. El coste puede medirse como el tiempo de retardo del enlace o, simplemente, el número de saltos entre redes hasta alcanzar el destino.

Y volvemos a enfrentarnos con el

problema de la cantidad de ordenadores que hay en Internet. Ningún *router* podría guardar un registro de cada ordenador, las distintas rutas para alcanzarlo y su coste asociado. En su lugar, como las redes están jerarquizadas, hay *routers* que dan conexión a subredes enteras, en las que todos los ordenadores tienen una parte de su dirección numérica igual (por ejemplo, todas empiezan por 80.57). De esta forma un único registro en la *tabla de enrutamiento*

sirve para describir el camino a miles o millones de ordenadores.

Diccionario de Términos

URL: dirección web que escribimos en el navegador.

Dominio: parte de la URL que identifica una organización.

IP: dirección numérica que entienden los ordenadores.

DNS: servidor de nombres capaz de traducir una URL en una IP.

DNS Raíz: servidor de nombres principal.

Enrutamiento: proceso para encontrar el camino hasta un ordenador.

IPv4: sistema actual de direcciones IP, compuesto de cuatro números hexadecimales.

IPv6: nuevo sistema de direcciones IP compuesto de seis números hexadecimales.

El Agotamiento de Direcciones

Las direcciones numéricas son asignadas por ICANN, un organismo sin ánimo de lucro. Y aunque con el sistema actual sistema de cuatro números por dirección hay más de 4.200 millones de direcciones, resulta insuficiente para todos los ordenadores, PDAs, coches, teléfonos, etc. que pueden llegar a utilizarse. Ya se ha puesto en marcha un nuevo sistema de direcciones numéricas, llamado IPv6 (el actual se llama IPv4), que permite 2^{128} direcciones. Eso son 340

sexillones de direcciones. O, dicho de otra forma, 670 mil billones de direcciones por cada milímetro cuadrado del planeta. Probablemente más que suficiente para cualquier eventualidad en los próximos siglos.

Y lo que quizás resulte más sorprendente, los mismos sistemas de búsqueda que empleamos hoy para los “escasos” 500 millones de servidores esperamos que sigan funcionando en el futuro. Y es que en ocasiones, la mejor solución es la más sencilla.

Bibliografía

- Wikipedia: <http://es.wikipedia.org>
- RFC 791 – Information Science Institute.
- Castells, M.: La Galaxia Internet – Reflexiones sobre Internet, empresa y sociedad. Barcelona (Plaza & Janés) 2001
- Tanenbaum, A.: Computer Networks Londres (Prentice Hall) 2002.