

Content Delivery Network (CDN)

"Content Delivery Network" (CDN) o Red de Entrega de Contenidos es la forma moderna de garantizar que nuestra página web pueda ser visitada por millones de usuarios y consiste, de manera muy simplificada, de repartir nuestro contenido estático a través de servidores ubicados físicamente en distintas partes del planeta, de tal manera que según sea la dirección IP del visitante se contacte con el servidor más cercano. Veamos ahora los detalles ¡vamos!

"Content Delivery Network" (CDN).

Para poder entender hacia dónde vamos debemos ver de dónde venimos.

El primer servidor web.

Tim Berner-Lee, inventor del hipertexto y del internet tal como lo conocemos, arrancó con un ordenador avanzado para la época fabricado por la compañía "NEXT Computers" la cual fue fundada por el legendario Steve Jobs luego de que lo despidieron de "Apple Computer", la empresa por él creada. En aquellos años 90 nosotros comprábamos la revista PCMagazine en castellano donde daban cuenta en un artículo que esas computadoras NEXT tenían la increíble cantidad de 1024 megabytes en RAM y costaba cada una más de 10 mil US\$ ¡sorprendente!

A dicho computador le pegaron un papelito indicando que era un servidor y que no debía ser apagado, estaba alojando la primera página web del mundo y pocos usuarios se conectaban a ella, específicamente los investigadores del CERN ("European Organization for Nuclear Research" Organización Europea para la Investigación Nuclear).

Montando nuestra primera página web.

Pues hoy en día tenemos los mismos recursos -incluso rebasan- que contaban y disponían los investigadores europeos: ordenadores con más de 8 mil megabytes de RAM y 4 núcleos de 64 bits con conexiones fijas a internet con más de 1 mbps de velocidad y sin ningún límite de megabytes son comunes, ***¿Qué nos impide montar en nuestro hogar u oficina nuestra página web?*** La limitación es que nuestro proveedor de internet muy probablemente nos proporcionará una dirección IP dinámica: al apagar el modem y volverlo a encender tendremos una dirección IP diferente y a nuestros usuarios tendremos que volver a notificarles del cambio. *Para evitar dicho inconveniente* desde los albores del nacimiento del internet se utilizan los DNS "Domain Name Service" Servicio de Nombres de Dominios: son ordenadores que uno les pregunta por un dominio (dirección web) y nos devuelve la dirección IP para conectarnos al ordenador que contiene la página web deseada.

Lo anterior lo podemos simplificar con los teléfonos móviles celulares: cuando una amiga os da su número simplemente lo ingresamos al teléfono y le damos marcar y listo, la encontramos y conectamos. Pero luego recordar ese número es el problema, para ello nuestro celular tiene una agenda donde podemos colocar el nombre de la chica asociada a su número telefónico. **En este ejemplo el DNS sería la agenda de contactos y el número de teléfono sería la dirección IP.** Huelga decir que los DNS son más avanzados que la agenda de contactos: si un dominio cambia su dirección IP ésta es automáticamente actualizada, cosa que no hace el programita de agenda de contactos del celular (a menos que llamara todos los días a la chica confirmando el número telefónico ¡*Qué locura!*).

Tráfico hacia nuestra página web.

Con este escenario, un ordenador en nuestra casa u oficina, con un acceso a internet las 24 horas del día sin límites de megabytes y una dirección IP fija y un DNS configurado estamos prestos a difundir al mundo nuestro trabajo sobre el software libre. Pero para hacerle honor a la verdad, lo que distingue un **servidor web** de un ordenador normal es la redundancia de sus componentes, así que tienen al menos dos fuentes de poder (si una se "quema" queda la otra funcionando y se puede cambiar *sin necesidad de apagar el servidor*), tienen al menos dos discos duros en RAID (escriben y leen los mismos datos en ambos discos) y también se pueden reemplazar sin apagarlo (al conectar un disco nuevo automáticamente la tarjeta madre copiará los datos del disco viejo al nuevo y "emparejará" la información), memoria RAM de seguridad que -adivinen- internamente son dobles y almacenan y comparan ambas copias constantemente **y también tienen dos o más tarjetas de red, generalmente ethernet con cables de cobre o fibra óptica, de igual manera si una falla queda la otra funcionando -y no se pueden reemplazar porque son integradas a la tarjeta madre pero se puede adicionar una(s) en una(s) ranura(s) PCI libre-**.

En cuanto a qué podremos hacer con un ordenador con dos o más tarjetas de red (si tenéis un ordenador portátil con tarjeta de red inalámbrica y ethernet podéis hacer pruebas) varios usos le podemos dar:

- Ambas tarjetas de red conectados a un concentrador/enrutador en una red de área local con diferentes direcciones IP (por supuesto de área local): podemos balancear la carga para enviar nuestra página web por ambas tarjetas, eso se configura fácilmente con **ifenslave** para luego aplicar **modprobe**.
- Una tarjeta de red conectada al modem y la otra al concentrador/enrutador para trabajarlo como "puente".
- Una tarjeta de red conectada al enrutador/concentrador y la otra a otro ordenador con una base de datos para aislar y aumentar la seguridad de nuestra información (páginas web dinámicas).
- Una tarjeta de red conectada a un concentrador y la otra a otro para crear dos segmentos de red.
- Muchas otras combinaciones (aquí imaginen ustedes).

Esto que acabamos de explicar es para ir "abriendo la mente", para lo que luego queremos explicar acerca de los CDN's.

Para continuar optaremos por configurar nuestro modem/concentrador/enrutador (hoy en día se puede dar los tres aparatos en uno, por eso decimos que tenemos más poder de hardware que los investigadores del CERN en los años 90). En todo caso explicamos que:

- Un modem, generalmente ADSL, nos conecta a internet.
- Un enrutador toma una dirección IP asignada a nuestro modem (que está en "modo puente") y traslada las peticiones (NAT) a nuestra red de área local (LAN).
- Un concentrador amplía la cantidad de puertos necesarios para conectar hasta 255 ordenadores en un segmento de red (la mayoría de enrutadores tienen 4 puertos físico alámbricos, son pocos en realidad).

Así las cosas nuestro enrutador recibe las peticiones y las reenvía a nuestro servidor web: hasta allí todo va bien, pero ¿qué sucede si recibimos más visitas?

Ampliando la cantidad de ordenadores en una LAN.

Con 4 ordenadores podemos realizar el siguiente esquema para atender más visitantes:

- Ordenador 1: recibe las peticiones y balancea la carga de trabajo hacia los ordenadores 2 y 3.
- Ordenadores 2 y 3 son los servidores web como tal, reciben las cabeceras de peticiones de ordenador 1.
- Ordenador 4: será donde guardemos nuestra página web y que los ordenadores 2 y 3 se encargarán de verificar constantemente si tienen la última versión de nuestra página web.

Este esquema tiene como ventaja, por supuesto, el poder atender más visitantes, manteniendo nuestro trabajo original en el ordenador 4 y teniendo respaldos en los ordenadores 2 y 3, **el punto débil de la cadena sería el ordenador 1, que balancea la carga ya que si falla se cae por completo nuestro dominio, ojo con eso.** Como desventaja podríamos mencionar, entre otras, el registro de los visitantes (quién ha sido atendido por ordenador 2 y quién por ordenador 3), problemas con las carpetas temporales en el caso que nuestros usuarios necesiten "subir" archivos a nuestra página web y con respecto al lenguaje PHP el almacenamiento de las variables de sesión que no serán iguales para todas las páginas (aquí debemos aclarar que nuestra página web es **un sitio web** conformado por varias páginas que pueden ser servidas tanto por ordenador 2 y ordenador 3: un visitante puede ver una hoja en ordenador 1 y al cabo del tiempo al volver a visitar el balanceo de carga lo envía a ordenador 3 *por lo tanto las variables de sesión PHP serán con distintos valores, lo que pudiera ocasionar problemas de programación -"amnesia" de las aplicaciones-*).

Añadiendo más "ancho de banda" a nuestra conexión a internet.

Pues que llega un momento que nuestro humilde modem ADSL no da para más velocidad, así que le pedimos a nuestro Proveedor de Servicio de Internet ("Internet Service Provider" ISP) nos suba la velocidad pero va a ser que no, lo máximo ya lo tenemos, 10 mbps y ADSL2 que llega a 25 mbps no está disponible... Pues agregamos otra línea telefónica y otro modem y **en teoría nuestro ISP debería colocarlos en modo vinculado "bonding": con la misma dirección IP transmite por ambos modem entre nuestra casa u oficina hasta la central telefónica más cercana.** ¿Recuerdan la redundancia de tarjetas de red en los servidores? Pues bueno, aquí es lo mismo, es una habilitación en una de las capas del modelo OSI, específicamente la capa física (de hecho cada modem en sí mismo cumple con esa capa y cada capa es independiente de las otras capas tanto por encima como por debajo).

Servidores espejos "mirrors".

Pues he aquí que nuestro negocio o portal educativo ha crecido y expandido en todo nuestro país, o incluso hacia el exterior , así que decidimos contratar alojamiento web en nuestro país y/o el exterior. En esos ordenadores contratados colocaremos una sincronización de archivos para que al modificar nuestros ficheros se copien, al cabo de cierto tiempo, a los servidores espejo.

Como ya tenemos configurado nuestro balanceo de carga lo único que tenemos que configurar es que se redirija hacia unas direcciones IP *que no están en nuestra red de área local, sino en internet.* **El pequeño detalle de esta configuración estriba en que nuestro alojamiento debe ser dedicado, tal como tenemos montado en nuestra casa u oficina, si es alojamiento compartido deberemos comprar otro dominio web porque de esa manera es que sabe un servidor web compartido qué datos mostrar al visitante.**

Si contratamos un servidor dedicado bien puede ser una máquina real o, lo que es más común hoy en día, una máquina virtual *que igual compartiremos con otros sitios web un ordenador físico pero con varias máquinas virtuales que tendrán sus propias tarjetas de red virtuales con sus propios y únicos MAC por los cuales se podrán identificar y asignar sus propias direcciones IP fijas (esto es particularmente atractivo para cuando migremos, por fin, a las direcciones IPv6)*

Inconvenientes de los servidores espejos.

Pues como ya se habrán percatado, nuestra debilidad sigue siendo el "balanceador de carga": nuestro dominio y por medio de los DNS sigue devolviendo una única dirección IP, la de nuestro ordenador en nuestra casa u oficina, *¡pues he aquí la brillante solución de los CDN!*

CDN "Content Delivery Network".

De nuevo simplificaremos al máximo para ilustrar de manera didáctica sobre el como funciona. A nivel mundial las direcciones IP están repartidas por grandes proveedores de acceso a internet o incluso por países enteros. Acá en Venezuela CANTV tiene su rango de direcciones IP asignadas, así que para los servidores web es "facil" determinar si nuestro visitante esta conectado por medio de ese proveedor de internet, simplemente comparando la dirección IP contra los rango registrados. Pueden incluso llegar a determinar desde cual ciudad se conectan. Basado en esto:

- Una empresa decide montar un CDN a nivel mundial, así que va país por país donde montar una granja de servidores con acceso a internet por medio de algún proveedor que a su vez tiene su rango de direcciones IP asignadas. Aparte de la compra de conexión en sí como tal, también se incluye la compra de un rango de direcciones IP para asignarlas a los ordenadores de la nueva granja.
- El siguiente paso es configurar con algunas de esas máquinas unos servidores DNS (generalmente uno principal y otro secundario). Dichos servidores DNS estarán a la orden de los servidores DNS del proveedor de internet y llevarán registro de los dominios web que van a alojar.
- Este procedimiento se repite país por país, o región por región.
- Una vez tengan "armada" la red es hora de ofrecer el servicio, no sin antes "unificar" todos sus servidores DNS para que operen como un conjunto armónico.

¿Cómo funciona el CDN?

Al comprar nosotros un alojamiento en CDN lo primero que tenemos que hacer es configurar ante quien nos vendió el nombre o dominio web y establecer los DNS de la empresa que ofrece el servicio CDN. Nuestro trabajo es seguir manteniendo nuestro "balanceador de carga" como siempre lo hemos hecho pero con privilegios especiales para que los ordenadores del CDN tengan acceso y copien nuestra página web y la distribuyan por cada una de las granjas de servidores que nombramos anteriormente.

Lo brillante de la idea de los CDN es analizar la dirección IP de un visitante de nuestra página web RECORDAD que primero tiene que pasar por los DNS del CDN para saber la dirección IP de nuestro balanceador de carga... PUES VA A SER QUE NO porque si nuestro visitante está más cerca de una granja de servidores del CDN pues le pasa esa dirección IP al visitante de una manera totalmente transparente, al fin y al cabo es una copia exacta de nuestra web lo que le va a ser retribuida al internauta.

Un asunto importante para un CDN es el poder identificar según la dirección IP recibida del webnauta de dónde viene para redirigirlo a la granja de servidores más cercana -y que contiene copia de nuestra página(s) web-. Es por ello que muchas empresas ofrecen para descargar sus bases de datos, con su respectivo "manual de uso", por ejemplo, desde **MaxMind** podeis visitar su página designada [al respecto en este enlace](#). Esencialmente se resume en este comando, siempre respetando y considerando *de no abusar del servicio*:

KS7000+WP

KS7000 migra a GNU/Linux y escoge a WordPress para registrar el camino.

<https://www.ks7000.net.ve>

```
wget http://geolite.maxmind.com/download/geoip/database/GeoLiteCity.dat.gz
z wget http://geolite.maxmind.com/download/geoip/database/GeoLiteCityv6-beta/GeoLiteCityv6.dat.gz
```

Dicha base de datos es utilizada [por aplicaciones de seguridad para proteger servidores web de ataques maliciosos](#). Dado el caso que suframos un ataque masivo de denegación de servicio (millones de dispositivos al mismo tiempo visitan nuestra página web con la intención de saturarla y colapsarla) pues simplemente será repartida a lo largo y ancho del planeta. *Puede darse el caso que el ataque provenga, por ejemplo, de millones de dispositivos infectados que están localizados físicamente en Europa: pues bien las granjas de servidores ubicadas en Europa (digamos que hayan tres, una en Francia, otra en Inglaterra y otra en Italia) **solo esas tres granjas de servidores serán colapsados, si es que pueden, mientras que las demás granjas en el resto del mundo mantendrán nuestra página web incólume.***

<https://www.youtube.com/watch?v=lnhGahmVjdk>

Conclusión.

Si queremos simplificar a tope podríamos decir, en una gran perspectiva, que los CDN actúan como unos "balanceadores de carga" pero a una gran medida, no sin antes realizar un arduo trabajo de sincronización de DNS y contenidos web. Esperamos les haya sido útil la información.

Fuentes consultadas.

En idioma castellano.

- [«CDN OVH: el complemento indispensable para sus sitios web internacionales»](#).

En idioma inglés.

- [«Apache, Load Balancers and Log Files»](#).
- [«Configure Apache Web Server Load Balancing»](#).

KS7000+WP

KS7000 migra a GNU/Linux y escoge a WordPress para registrar el camino.
<https://www.ks7000.net.ve>

- «[Load balancing \(computing\)](#)».
- «[CDN - Content Delivery Network](#)» by Vangie Beal.
- «[Mirror site](#)» by Vangie Beal.
- «[What Is a Content Delivery Network?](#)».
- «[Bonded ADSL vs Traditional ADSL](#)».
- «[Content Distribution Network](#)» by Akamai.
- «[GeoIP Legacy Downloadable Databases](#)» by MaxMind.
- «[How to Install Bro on Ubuntu 16.04](#)» by Finide at DigitalOcean.

Recursos multimedia.

- «[File:NCDN - CDN.png](#)»
- «[CDN Explained in 60 Seconds - Content Delivery Network](#)»